

Sequence Match Listing for  
09/964824

for SETA ID NO: 137

RESULT 2  
AB011095/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

AB011095  
Homo sapiens mRNA for KIAA0523 protein, partial cds.  
AB011095  
KIAA0523 protein.  
Homo sapiens male brain cDNA to mRNA, clone\_11b:pb1uescriptII SK plus clone:HG1394.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.  
Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro  
DNA Res. 5 (1), 31-39 (1998)  
98390545  
2 (bases 1 to 5172)  
Ohara, O., Nagase, T. and Ishikawa, K.  
Direct Submission  
Submitted (13-FEB-1998) Otsamu Ohara, Kazusa DNA Research Institute, DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:cdna1info@kazusa.or.jp, Tel: +81-438-52-3913, Fax: +81-438-52-3914)  
Location/Qualifiers  
1. 5172  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HG1394"  
/sex="male"  
/tissue\_type="brain"  
/clone\_11b="pb1uescriptII SK plus"  
/gene="KIAA0523"  
c1. 1407

/gene="KIAA0523"  
/codon\_start=1  
/product="KIAA0523 protein"  
/protein\_id="BA25449.1"  
/db\_xref="GI:3043570"  
/translation="ELRLRRRMFHHPMSDSQPPALGPAPPAIRHSRGTYICFSD  
DGHRTLKGAVFYDLKMTVSHCODACAESESYAGAEABECYGRNLRAVSGLBE  
CNHECKGKSGVAGVDRLSYVRDELQPSRRKRIATYIGCTRPENITHAPSSLI  
QAMVTGTCGSGFSOKREFFAIIIRGECICATIPPRNLADAMDSYGGDDPEAKRLA  
FDCTLYIKGFGKGDHMRSRRTICVTHESGREIIEFDGSAIILIRPYSILVAEPNR  
KAGHLGTAADRMWKSKEWDFVNSYASMSHLDLKYGRLLVYHYBELRSLVP  
TLREWVAFLNVSSEERLLCVENNKESFRRRRRSHDEPFPPEMDLINGYIRFVVD  
QALRDHNMFTGLPREYVPR"  
BASE COUNT 1191 a 1340 c 1492 g 1149 t  
ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 5172;  
Best Local Similarity 100.0%; Pred. No. 2, 4e-68;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACACATTAACAGAAATTTTACCTCATATACCAAAAAAGTCACAGAGTAAAGCCAAAC 60  
DB 5172 AACACATTAACAGAAATTTTACCTCATATACCAAAAAAGTCACAGAGTAAAGCCAAAC 60  
QY 61 TCAGCAGAGCTTGATCTCTGCTACTTAACAAATTTACCAAGACCTGATCTCTTCTGCC 120  
DB 5112 TCAGCAGAGCTTGATCTCTGCTACTTAACAAATTTACCAAGACCTGATCTCTTCTGCC 120  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGCTGATCTCTGCTATGATCCCAAGGCC 180  
DB 5052 TCTCAACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGCTGATCTCTGCTATGATCCCAAGGCC 180  
QY 181 CAAGTGTGCTATTAACAGCCAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
DB 4992 CAAGTGTGCTATTAACAGCCAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
QY 241 TTAACAGCTTCTACCCAGCAT 263  
DB 4932 TTAACAGCTTCTACCCAGCAT 4910

RESULT 3  
AC007962/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

AC007962  
Homo sapiens chromosome 17 clone 2511\_J\_5 map 17, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
AC007962  
HTG: HTGS\_PHASE1.  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 172091)  
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Baran, N., Beckley, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, M., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., Jones, C., Kahn, L., Karas, A., Lehoczy, J., Lien, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Molla, M., Morris, M., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., O'Connor, T., O'Donnell, P., Pavlin, B., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Testaye, S., Tortuella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,

23 10:00:33 2003

us-09-964-824c-137.rge

Page 3

## TITLE

Wheeler, J., Wu, X., Myman, D., Ye, W. J. and Zody, M.  
Direct Submission

## JOURNAL

Submitted (03-JUL-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker: Smit, A.F.A. &  
Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

## COMMENT

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 25 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2475: contig of 2475 bp in length  
2476 4789: gap of unknown length  
4790 7000: gap of unknown length  
7001 9027: gap of unknown length  
9028 11768: gap of unknown length  
11769 15071: gap of unknown length  
15072 19326: gap of unknown length  
19327 22239: gap of unknown length  
22240 26006: gap of unknown length  
26007 30364: gap of unknown length  
30365 34450: gap of unknown length  
34451 38286: gap of unknown length  
38287 42572: gap of unknown length  
42573 47699: gap of unknown length  
47700 54016: gap of unknown length  
54017 60547: gap of unknown length  
60548 68280: gap of unknown length  
68281 75214: gap of unknown length  
75215 80915: gap of unknown length  
80916 87222: gap of unknown length  
87223 94562: gap of unknown length  
94563 106577: gap of unknown length  
106578 119209: gap of unknown length  
119210 141857: gap of unknown length  
141858 172091: gap of unknown length

## FEATURES

## source

1. 172091  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="2511\_J-5"  
/clone.lib="CITB Human BAC"  
BASE COUNT 44773 a 39840 c 42000 g 44968 t 510 others

## ORIGIN

Query Match 100.0%; Score 263; DB 2; Length 172091;  
Best Local Similarity 100.0%; Pred. No. 3.9e-66;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AACCATATAACAGATTTATATAGTCATATATACAAAAAAGTCAGAGGTAAAGCCATC 60  
DB 80393 AACCATATAACAGATTTATATAGTCATATATACAAAAAAGTCAGAGGTAAAGCCATC 80334  
QY 61 TCAAGCAGGCTTGATCTGTAATCTTAAACATTTACAGAGGACTGATCTTTCTGCC 120  
DB 80333 TCAAGCAGGCTTGATCTGTAATCTTAAACATTTACAGAGGACTGATCTTTCTGCC 80274  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAGCTTCAGGTGATTCGTGTCATGATCCAGAGCC 180  
DB 80273 TCTCAACTCTCCCTTCAGTGTGTCAGCTTCAGGTGATTCGTGTCATGATCCAGAGCC 80214  
QY 181 CAAGGTGTCATCATTAAGACCCAGAGATCTACTACTTTTTCACATTTCAACAGGGAA 240  
DB 80213 CAAGGTGTCATCATTAAGACCCAGAGATCTACTACTTTTTCACATTTCAACAGGGAA 80154  
QY 241 TTAACACAGCTTCTACCCAGCAT 263  
DB 80153 TTAACACAGCTTCTACCCAGCAT 80131

## ALIGNMENTS

RESULT 1

ABL67103

ID ABL67103 standard; DNA; 263 BP.

XX

AC ABL67103;

XX

DT 15-MAY-2002 (first entry)

XX

DE Thyroid cancer related gene sequence SEQ ID NO:5440.

XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

PF 30-MAY-2001; 2001WO-US10838.

XX

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.

for SEQ ID NO. 137

Seq ID NO: 137 - cont -

PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 28-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 29-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 02-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.

## (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horigan S;  
PI Soppet DR, Weaver Z;

WP1: 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
XX determining a change in expression of a gene of a signature gene set -

PS Claim 1; SEQ ID 5440; 44pp; English.

XX The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABU61664  
CC to ABU70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.

SQ Sequence 263 BP; 85 A; 67 C; 42 G; 69 T; 0 other;

Query Match 100.0%; Score 263; DB 24; Length 263;  
Best Local Similarity 100.0%; Pred. No. 2e-71;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAACAATTAACAGATTATAGCTATATACAAAAAAGTCCAGAGTAAGGCCAATC 60  
|||||

Db 1 AAACAATTAACAGATTATATAGCTATATACAAAAAAGTCCAGAGTAAGGCCAATC 60  
OY 61 TCAAGCCAGGCTTGATCTCTGTAATTAACAAATTTACCAAGAGCTTGATCTCTCTGCGC 120  
Db 61 TCAAGCCAGGCTTGATCTCTGTAATTAACAAATTTACCAAGAGCTTGATCTCTCTGCGC 120  
OY 121 TCTCAACTCTCCCTCAGAGGTGTCAGCTTCACGTGATCTCTGTCATGATCCCAAGGCC 180  
Db 121 TCTCAACTCTCCCTCAGAGGTGTCAGCTTCACGTGATCTCTGTCATGATCCCAAGGCC 180  
OY 181 CAAGGTGTCATCATTAAGACCAGGAATACCTTTTTCACATTCACAGAGGAA 240  
Db 181 CAAGGTGTCATCATTAAGACCAGGAATACCTTTTTCACATTCACAGAGGAA 240  
OY 241 TTAACACAGCTTCTACCCAGCAT 263  
Db 241 TTAACACAGCTTCTACCCAGCAT 263

New sensitive

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 20:42:18 ; Search time 779.199 Seconds

(without alignments)  
5466.401 Million cell updates/sec

Title: US-09-964-824C-137

Perfect score: 263  
Sequence: 1 aaacaataacagacattat.....aaacagcttaccacagcat 263Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estlin: \*  
4: em\_estlun: \*  
5: em\_estlov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hlc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hlc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estlom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_iny: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_rtd: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	263	100.0	263	9	AA426220 zv84d01.s
2	263	100.0	300	9	A1915364 A1915364
3	263	100.0	309	14	BM930947 UT-E-EJ0-BM930947
4	263	100.0	312	14	BM684449 UT-E-EJ0-BM684449
5	263	100.0	348	14	Z39512 HSC1BH062
6	263	100.0	352	14	N63086 yz32e07.s1

7	263	100.0	397	9	A1095108	A1095108 ga14h11.x
8	263	100.0	440 <td>9</td> <td>A1056703</td> <td>A1056703 oys3e04.x</td>	9	A1056703	A1056703 oys3e04.x
9	263	100.0	455	10	BE326640	BE326640 h62f06.x
10	263	100.0	470	9	A1359305	A1359305 qy29c09.x
11	263	100.0	481	9	A1421275	A1421275 t119g05.x
12	263	100.0	491	10	AW008726	AW008726 ws72c12.x
13	263	100.0	497	9	A1685123	A1685123 wc67g02.x
14	263	100.0	512	10	AW294233	AW294233 ut-H-B12-BM676160
15	263	100.0	522	14	BM676160	BM676160 ut-E-EJ0-AA129740
16	263	100.0	540	9	AA129740	AA129740 zpn1e08.s
17	263	100.0	541	9	A1628620	A1628620 ly77c03.x
18	263	100.0	547	10	BE504506	BE504506 h257d03.x
19	263	100.0	564	14	BM682221	BM682221 ut-E-EJ0-AW290920
20	263	100.0	588	10	AW290920	AW290920 ut-H-B12-F02819
21	262	99.6	270	14	F02819	AA938912 op74f09.s
22	261.4	99.4	583	9	AA938912	A1916307 we29c10.x
23	261.4	99.4	583	9	A1916307	A1191041 qe25a07.x
24	258.2	98.2	272	9	A1191041	H29606 ym61d05.s1
25	257.8	98.0	509	14	H29606	H29606 yd99d03.s1
26	256.8	97.6	317	14	T89670	AA992195 ot80e07.s
27	251	95.4	405	9	AA992195	BM723364 ut-E-EJ0-H08230
28	251	95.4	666	14	BM723364	BM723364 y192g05.s1
29	247	93.9	447	14	H08230	H24358 ym56b01.s1
30	246	93.5	432	14	H24358	T91053 yd59c04.s1
31	245	93.2	413	14	T91053	A1868809 tcs1904.x
32	243	92.4	262	9	A1868809	AQ0701917 HS-5396.B
33	240.6	91.5	548	17	AQ0701917	A186979 ut-H-B11-BG170361
34	236.4	89.9	374	10	AM136979	H66033 yr70b02.s1
35	234.8	89.3	512	12	BG170361	H23292 ym57b09.s1
36	233.4	88.7	417	14	H66033	T31702 EST37298.Hu
37	228.4	86.8	420	14	H23292	BG108592 602278574
38	204	77.6	234	14	T31702	H18630 ym45a08.s1
39	186.8	71.0	761	12	BG108592	BF155474 OY4-HR053
40	178.6	67.9	392	14	H18630	BM712977 ut-E-EJ0-A1609882
41	150.8	57.3	448	12	BF155474	A1610939 t160e06.x
42	144	54.8	375	14	BM712977	BF944654 PM0-NN117
43	139.2	52.9	194	9	A1609882	
44	138.2	52.5	181	9	A1610939	
45	113.4	43.1	655	12	BF944654	

## ALIGNMENTS

RESULT 1  
AA426220 263 bp mRNA linear EST 16-OCT-1997  
DEFINITION zv84d01.s1 Soares-total-fetus-ND2HF8\_9w Homo sapiens cDNA clone  
IMAGE:760321 3', mRNA sequence.  
AA426220  
AA426220.1 GI:2107753

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

human.

Homo sapiens

REFERENCE  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 263)AUTHORS  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellander, K., Stepien, K., Tan, F., Theisling, B., White, T., Wyllie, T., Waterston, R. and Wilson, R.WashU-Merck EST Project 1997  
Unpublished (1997)  
Contact: Wilson, R.K.TITLE  
JOURNAL  
COMMENTWashington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810Email: est@wustl.edu  
This clone is available royalty-free through LINT; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.Seq primer: -41ml3 fwd. RT from Amersham  
High quality sequence stop: 194.

## FEATURES

## Source

Location/Qualifiers  
1. .263

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

```
/db_xref="taxon:9606"
```

```
clone_lib="Soares_total_fetus_Nb2HF8_9w"
```

```
/lab_host="DH10B"
/notes="Vector: pTZ19-Dac (DharmaSci)" with
```

```
/lab_host="DH10
```

**Note:** Vector: p773-Pac (Pharmacia) with a modified polylinker: Site 1: Not I, Site 2: Eco RI, 1st strand cDNA (total) fetus material with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAACGAGGCGCGCGCTATTTTCTTTTCTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by

BASE COUNT	85 a	67 c	42 g	69 t
------------	------	------	------	------

Query Match  
Best Local

100.08; Score 263; DB 9; Length 263;  
100.08; Prod No 4 30-61.

```

100.00%, FID. NO. 4.3E+01,
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

QY	1	AAACAAATTAACAGAAATTTATATAGTCATATTAACAAAAAAGTCAGAGGTAAGGCCAATC	60
Db	1	AAACAAATTAACAGAAATTTATATAGTCATATTAACAAAAAAGTCAGAGGTAAGGCCAATC	60
QY	61	TCAAGCAAGGCTTGATCCTGTACTTAAACAAATTCACCAGAGACTGTACTCTTCTTGCC	120
Db	61	TCAAGCAAGGCTTGATCCTGTACTTAAACAAATTCACCAGAGACTGTACTCTTCTTGCC	120
QY	121	TCTCAACTCTCCCTTCAGTGGTGTACAGTTTCACGATTCCTGGTCATGATCTCCAAAGCC	180
Db	121	TCTCAACTCTCCCTTCAGTGGTGTACAGTTTCACGATTCCTGGTCATGATCTCCAAAGCC	180
QY	181	CAAGTGGTCAATCAATAAGACCCAGAACTACTACTACCTTTTTCACATTCACAGAGGGA	240
Db	181	CAAGTGGTCAATCAATAAGACCCAGAACTACTACTACTCTTTTTCACATTCACAGAGGGA	240
QY	241	TTTAAACAGCTTTCACCCAGCAT	263
Db	241	TTTAAACAGCTTTCACCCAGCAT	263

RESULT 2  
A1915364  
00015364  
00015364

LOCUS	3381204	300 bp	mRNA	linear	EST
DEFINITION	w38108.x1 Soares_MFL_T-GSC_S1 Homo sapiens CDNA clone				
ACCESSION	IMAGE:2330439 3', mRNA sequence.				
VERSION	AI915364				
KEYWORDS	AI915364.1 GI:5635219				
SOURCE	EST.				
	human.				

```

/db_xref="taxon::9606"
/clone="IMAGE:2330439"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"

```

```
/lab_host="DH10B"
```

a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;

libraries (fetal lung NBHL19W, testis NHT, and B-cell

NI: CGAP, CGB1) were mixed, and ss circles were made in PCR. Following HAP purification, this DNA was used as tracer in a subtractive cDNA hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5 000 clones made from the same 3 libraries (the pools consisted of T.M.A.C.E. clones 2974680-302087, 682632-687239, 726408-726811, and 729036-731539, Subtraction by Bento Soares and M. Faíma Bonaldo. "

Matches 26

0; Mismatches 0; Indels

1 AAACAATAAACAGAAATTTATTAGCTCATATATACAAAAAAAGTCCAGAGGTAAAGGCCAATC 60

[illegible]

LOCUS	BM930947	309 bp	mRNA	linear	EST J3-MAR-200
DEFINITION	UI-E-EJ0-a1p-a-08-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone				

ACCESSION	U1E-EJO- <del>alp</del> -a-08-0-01 5', mRNA sequence.
VERSION	BM930947
KEYWORDS	BM930947 1 GI:19390120
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

JOURNAL  
COMMENT

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cqapbs-remail.nlh.gov](mailto:cqapbs-remail.nlh.gov)

LINE 10: This clone is available royalty-free through LNL ; contact the IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.

Insert Length: 459      Std Error: 0.000

Seq primer: -40UP from Gibco  
High quality sequence stop: 272

High quality sequence stop: 272

FEATURES	Location/Qualifiers
source	1. .300

/organism="Homo sapiens"

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA sequence: 276-305, >AT-rich#low-complexity (matched complement)  
Seq primer: M13 REVERSE.

FEATURES  
Source  
Location/Qualifiers

1..309  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-aip-a-08-0-UI"  
/clone\_lib="UI-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site.1: EcoR I; Site.2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCCAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT  
ORIGIN  
97 a 45 c 76 g 91 t

Query Match  
Best Local Similarity 100.0%; Score 263; DB 14; Length 309;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAGTCCAGAGTAAGCCCAATC 60  
Db 285 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAGTCCAGAGTAAGCCCAATC 226  
QY 61 TCAAGCAAGCTTGATCTGCTACTTAACATTTCCACCAAGACTTGATCTTTGCCC 120  
Db 225 TCAAGCAAGCTTGATCTGCTACTTAACATTTCCACCAAGACTTGATCTTTGCCC 166  
QY 121 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATGCTGCTGATGCTGATGCTGATGCTG 180  
Db 165 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATGCTGCTGATGCTGATGCTGATGCTG 106  
QY 181 CAAGTGGTATCATTAAGACCCAGATTAAGTACTTACATTTTACATTCACAGGGGAA 240  
Db 105 CAAGTGGTATCATTAAGACCCAGATTAAGTACTTACATTTTACATTCACAGGGGAA 46  
QY 241 TTTAAACAGCTTACCCAGCAT 263  
Db 45 TTTAAACAGCTTACCCAGCAT 23

RESULT 4  
BM684449

LOCUS  
DEFINITION  
UI-E-EJ0-aip-a-08-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone

ACCESSION  
VERSION  
KEYWORDS  
BM684449.1 GI:18994345  
EST.

SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 312)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL  
MEDLINE  
CONTACT: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA sequence: 1-28, >AT-rich#low-complexity (matched complement)  
Seq primer: M13 Forward  
POLYA-Yes.

FEATURES  
Source  
Location/Qualifiers

1..312  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-aip-a-08-0-UI"  
/clone\_lib="UI-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina foveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site.1: EcoR I; Site.2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCCAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG\_LIB=UI-E-EJ0  
TAG\_TISSUE=Foveal and Macular Retina  
TAG\_SEQ=GTCC"

BASE COUNT  
ORIGIN  
91 a 79 c 45 g 97 t

Query Match  
Best Local Similarity 100.0%; Score 263; DB 14; Length 312;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAGTCCAGAGTAAGCCCAATC 60  
Db 19 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAGTCCAGAGTAAGCCCAATC 78  
QY 61 TCAAGCAAGCTTGATCTGCTACTTAACATTTCCACCAAGACTTGATCTTTGCCC 120  
Db 79 TCAAGCAAGCTTGATCTGCTACTTAACATTTCCACCAAGACTTGATCTTTGCCC 138  
QY 121 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATGCTGCTGATGCTGATGCTGATGCTG 180  
Db 139 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATGCTGCTGATGCTGATGCTGATGCTG 198







Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo.  
BASE COUNT 134 a 110 c 80 g 116 t  
ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 440;  
Best Local Similarity 100.0%; Pred. No. 4,9e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 60  
DB 2 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 61  
OY 61 TCAAGCAGGCTTGATCTCTACTTAACCAATTTCACCAAGGACTGATCTTTGCG 120  
DB 62 TCAAGCAGGCTTGATCTCTACTTAACCAATTTCACCAAGGACTGATCTTTGCG 121  
OY 121 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATTCACCAAGGCC 180  
DB 122 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATTCACCAAGGCC 181  
OY 181 CAAGTGTGTCATATAAGACCAGCAATCTACTACTTTTTCACATTCACAGGGAA 240  
DB 182 CAAGTGTGTCATATAAGACCAGCAATCTACTACTTTTTCACATTCACAGGGAA 241  
OY 241 TTAACAAGCTTCTACCCAGCAT 263  
DB 242 TTAACAAGCTTCTACCCAGCAT 264

RESULT 9  
BE326640 455 bp mRNA linear EST 14-JUL-2000  
LOCUS br62f06.x1 NCL\_CGAP\_kid11 Homo sapiens cDNA IMAGE:3133091 3',  
DEFINITION mRNA sequence.  
ACCESSION BE326640  
VERSION BE326640.1 GI:9200416  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
COMMENT Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-f@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL, send email to:  
infoimage.lnl.gov  
Seq primer: -40UP from Gibco.

FEATURES  
source  
Location/Qualifiers  
1. 455

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3133091"  
/clone\_lib="NCI\_CGAP\_kid11"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site:1: Not I; Site:2: Eco RI;  
Plasmid DNA from the normalized library NCI\_CGAP\_kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneids 132376-132391, 1456007-145675, and

1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo.  
BASE COUNT 140 a 111 c 87 g 117 t  
ORIGIN

Query Match 100.0%; Score 263; DB 10; Length 455;  
Best Local Similarity 100.0%; Pred. No. 4,9e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 60  
DB 1 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 60  
OY 61 TCAAGCAGGCTTGATCTCTACTTAACCAATTTCACCAAGGACTGATCTTTGCG 120  
DB 62 TCAAGCAGGCTTGATCTCTACTTAACCAATTTCACCAAGGACTGATCTTTGCG 120  
OY 121 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATTCACCAAGGCC 180  
DB 122 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATTCACCAAGGCC 180  
OY 181 CAAGTGTGTCATATAAGACCAGCAATCTACTACTTTTTCACATTCACAGGGAA 240  
DB 182 CAAGTGTGTCATATAAGACCAGCAATCTACTACTTTTTCACATTCACAGGGAA 240  
OY 241 TTAACAAGCTTCTACCCAGCAT 263  
DB 242 TTAACAAGCTTCTACCCAGCAT 263

RESULT 10  
A1359305 470 bp mRNA linear EST 15-FEB-1999  
LOCUS qy29c09.x1 NCI\_CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:2013424 3',  
DEFINITION mRNA sequence.  
ACCESSION A1359305  
VERSION A1359305.1 GI:4110926  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
TITLE NCI/MINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
JOURNAL National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-f@mail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
www-bio.lnl.gov/brp/image/image.html  
Insert Length: 730 Std Error: 0.00  
Seq primer: -40UP from Gibco.

FEATURES  
source  
Location/Qualifiers  
1. 470

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2013424"  
/clone\_lib="NCI\_CGAP\_Brn23"  
/tissue\_type="q1idiolastoma (pooled)"  
/lab\_host="DH10B"  
/note="Organ: Brain; Vector: pT7T3D-Pac (Pharmacia) with a  
modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
GTGTACCAATCTGAAGTGGAGCGCGCATATCTTTTCTTTTCTTTTCTTTT

T 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. The library is normalized, and was constructed by Benton and Davis [1977].

BASE COUNT	145 a	111 c	94 g	120 t
ORIGIN				

ORIG

Query Match	100.0%;	Score 263;	DB 9;	Length 470;
Best Local Similarity	100.0%;	Pred. No. 4.9e-61;		
Matches 263; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

OY	1	AAACATTAACGGAATTATTTAGTCATATAACAAAAAGTCCAGGTAAAGGCCAATC	60
Dd	1	AACCATTAACGGAATTTATTTAGCTCATATAACAAAAAAGTCCAGGGAAGGCCAATC	60
OY	61	TCAAGCAAGGCTTGATCCTGTACTTAACAATTTCCACCAAGGACTTGATCTCTTCTGCC	120
Dd	61	TCAACCAAGGCTTGATCCTGTACTTAACAATTTCCACCAGGACTTGATCTCTTCTGCC	120
OY	121	TCTCAAGCTCCCTTCAGTGGTGCATCGCTTCACGTGATTCCTGTATATGCCAAGGCC	180
Dd	121	TCTCAAGCTCCCTTCAGTGGTGCATCGCTTCACGTGATTCCTGTATATGCCAAGGCC	180
OY	181	CAAGTGTGTCATCATTAAGACCCAGAGAATACACTTCTTTCACATTCAACAGGGGAA	240
Dd	181	CAAGTGTGTCATCATTAAGAACCCAGAGAATACACTTCTTTCACATTCAACAGGGGAA	240
OY	241	TTAAAACAGCTTCTACCCAGCAT	263
Dd	241	TTAAAACAGCTTCTACCCAGCAT	263

RESULT 11	
AI421276	
LOCUS	481 bp mRNA
DEFINITION	linear EST 28-MAR-1999
	U11905.X1 NCI_CGAP_Brn23 Homo sapiens CDNA IMAGE:2096696 3' ,
	mRNA sequence.

VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AI421276.1	GI:4267207	human.	Homo sapiens	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
			(bases 1 to 481)	

TITLE	National Cancer Institute/ National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BRGAP), Tumor Gene Index
JOURNAL	Unpublished (1998)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-femail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.

FEATURES	Location/Qualifiers
source	1. .481

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2096696"
/clone_lib="NCI_GCAP_Brn23"
/tissue_type="gastroblastoma (pooled)"
/lab_host="DH10B"

```

/note="Oran; Brain; Vector pT73D-Pa; Site:2; Eco RI; with a modified polylinker; site:1; Not I - oligo(dT) primer [5']  
strand cDNA was primed with a Not I - oligo(dT) primer [5']  
TGTCACCATGCAAGCGGAGCGGCCGCATACACTTTTTTTTTTTTTTTT  
T 3'); double stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT73 vector.  
Library is normalized, and was constructed by Bento  
Scares and M.Fátima Bonaldo.

BASE COUNT	149 a	114 c	96 g	122 t
ORIGIN				

ORIGIN

Query Match	100.0%;	Score 263;	DB 9;	Length 481;
Best Local Similarity	100.0%;	Pred. No. 5e-61;		
Matches 263;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	AAACATTAACGAAATTTATTAGCTCATATTAACAAAAAAGTCCAGAGGTAAAGGCCAATC	60
Db	1	AAACATTAACGAAATTTATTAGCTCATATTAACAAAAAAGTCCAGAGGTAAAGGCCAATC	60
QY	61	TCAACCAAGGCTTGATCCTGTACTTAAACATTTCCACCAAGGACTTGATCTCTTCTGCC	120
Db	61	TCAACCAAGGCTTGATCCTGTACTTAAACATTTCCACCAAGGACTTGATCTCTTCTGCC	120
QY	121	TCTCAACTCTCCCTTCAGTGGTGTAGCTTCACGATTCCTGGTATATATCCCAAGGCC	180
Db	121	TCTCAACTCTCCCTTCAGTGGTGTAGCTTCACGATTCCTGGTATATATCCCAAGGCC	180
QY	181	CAAGTGGTCAATATAAAGACCACAGAAATACTACTCTTTTTCACATTTCAACAGGGGAA	240
Db	181	CAAGTGGTCAATATAAAGACCACAGAAATACTACTCTTTTTCACATTTCAACAGGGGAA	240
QY	241	TTAAACAGCTTCTACCCAGCAT	263
Db	241	TTAAACAGCTTCTACCCAGCAT	263

RESULT 12	
AW008726	
LOCUS	AW008726
DEFINITION	491 bp mRNA linear EST 08-MAR-2000
	ws12c12.x1 NCI-CCAP_Brn23 Homo sapiens cDNA clone IMAGE:250272 3'
	mRNA sequence.

VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AM008726.1	GI:5857504	EST.	human. Homo sapiens	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
			(bases 1 to 491)	

**TITLE** National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CCAP/BTCAP), Tumor Gene Index  
**JOURNAL** Unpublished (1998)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
Email: cgaops-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld

FEATURES	Location/Qualifiers
source	1. .491

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

```
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="gliblastoma (pooled)"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCGAGTGGAGGCGGCGGATCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Farima Bernaldo."
```

BASE COUNT 153 a 118 c 96 g 123 t 1 others

ORIGIN

Query Match 100.0%; Score 263; DB 10; Length 491;  
Best Local Similarity 100.0%; Pred. No. 5e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCATTAACAGAAATTTATTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 60  
Db 1 AACCATTAACAGAAATTTATTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 60  
QY 61 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACAGAGAGCTGATCTTTCTGCC 120  
Db 61 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACAGAGAGCTGATCTTTCTGCC 120  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAAGCTTCAAGTATCTGTCATGATCCCAAGGCC 180  
Db 121 TCTCAACTCTCCCTTCAGTGTGTCAAGCTTCAAGTATCTGTCATGATCCCAAGGCC 180  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAAGCTTCAAGTATCTGTCATGATCCCAAGGCC 180  
Db 121 TCTCAACTCTCCCTTCAGTGTGTCAAGCTTCAAGTATCTGTCATGATCCCAAGGCC 180  
QY 181 CAAGGTGGTCATCATTAAGACCCAGAAATCTACTACTTTTCACATTAACAGGGGAA 240  
Db 181 CAAGGTGGTCATCATTAAGACCCAGAAATCTACTACTTTTCACATTAACAGGGGAA 240  
QY 241 TTAACACAGCTTCTACCCAGCAT 263  
Db 241 TTAACACAGCTTCTACCCAGCAT 263

RESULT 13  
A1685123 497 bp mRNA linear EST 17-DEC-1999  
LOCUS wc67602.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA IMAGE:2323730 3',  
DEFINITION mRNA sequence.  
ACCESSION A1685123  
VERSION A1685123.1 GI:4896417  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 497)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-remail.nih.gov  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Cloned distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbp/image/image.html  
Insert Length: 766 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 412.  
Location/Qualifiers  
1..497  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2323730"

```
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-Sport6; Site_1: SalI;  
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.72 Kb. Life Technologies catalog #:  
11548-013"
```

BASE COUNT 154 a 120 c 97 g 126 t

ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 497;  
Best Local Similarity 100.0%; Pred. No. 5e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCATTAACAGAAATTTATTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 60  
Db 2 AACCATTAACAGAAATTTATTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 61  
QY 61 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACAGAGAGCTGATCTTTCTGCC 120  
Db 62 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACAGAGAGCTGATCTTTCTGCC 121  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAAGCTTCAAGTATCTGTCATGATCCCAAGGCC 180  
Db 122 TCTCAACTCTCCCTTCAGTGTGTCAAGCTTCAAGTATCTGTCATGATCCCAAGGCC 181  
QY 181 CAAGGTGGTCATCATTAAGACCCAGAAATCTACTACTTTTCACATTAACAGGGGAA 240  
Db 182 CAAGGTGGTCATCATTAAGACCCAGAAATCTACTACTTTTCACATTAACAGGGGAA 241  
QY 241 TTAACACAGCTTCTACCCAGCAT 263  
Db 242 TTAACACAGCTTCTACCCAGCAT 264

RESULT 14  
AM294233 512 bp mRNA linear EST 16-JAN-2000  
LOCUS UT-H-B12-abn-f-12-0-UT.s1 NCI\_CGAP\_Sub4 Homo sapiens cDNA clone  
DEFINITION IMAGE:2726542 3', mRNA sequence.  
ACCESSION AM294233  
VERSION AM294233.1 GI:6700869  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 512)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-remail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbp/image/image.html The following repetitive  
elements were found in this cDNA sequence: 1-27,  
>AT-rich#Low-complexity  
Seq primer: M13 Forward  
POLYA=Yes.  
Location/Qualifiers  
1..512  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2726542"  
/clone\_lib="NCI\_CGAP\_Sub4"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker. Site\_1: Not I; Site\_2: Eco RI; The NCI\_CGAP\_Sub4 library is a subtracted library derived from the NCI\_CGAP\_Sub2 library which is a subtracted library derived from the NCI\_CGAP\_Sub1 library, which is a subtracted library derived from Bi. Bi constitutes a mixture of 21 normalized or subtracted NCI\_CGAP libraries: NCI\_CGAP\_C04, NCI\_CGAP\_P122, NCI\_CGAP\_P28, NCI\_CGAP\_C010, NCI\_CGAP\_C016, NCI\_CGAP\_K1d5, NCI\_CGAP\_K1d12, NCI\_CGAP\_K1d3, NCI\_CGAP\_C08, NCI\_CGAP\_P121, NCI\_CGAP\_Ly24, NCI\_CGAP\_Brn23, NCI\_CGAP\_Lu5, NCI\_CGAP\_C1L1, NCI\_CGAP\_Lu24, NCI\_CGAP\_Lu19, NCI\_CGAP\_C04, NCI\_CGAP\_C06, NCI\_CGAP\_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:

NCI\_CGAP\_K1d3 pool 1: LLM 3334-3337, 3682-3683, 3798-3803 (IMAGE cloneids 1322376-1323811, 1456008-1456775, 1500552-1502853) NCI\_CGAP\_K1d5 pool 1: LLM 3338-3342, 3722-3725, 3776-3778 (IMAGE cloneids 1333912-1325831, 1471368-1472903, 1492104-1493255) NCI\_CGAP\_Lu5 pool 1: LLM 3575-3582, 3851-3854 (IMAGE cloneids 1414920-1417991, 1520904-1522439) NCI\_CGAP\_C04 pool 1: LLM 3164-3167, 3716-3720, 3733-3735 (IMAGE cloneids 1257096-1258631, 1469064-1470993, 1475592-1476743) NCI\_CGAP\_P122 pool 1: LLM 2457-2459, 2758-2759, 3062-3068 (IMAGE cloneids 985608-986759, 1101192-1101959, 1217928-1220615) NCI\_CGAP\_C010 pool 1: LLM 2644-2653, 2871-2872 (IMAGE cloneids 1057416-1061255, 1144584-1145351) Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.] TAG\_LIB=NCI\_CGAP\_Brn23 TAG\_TISSUE=brain TAG\_SEQ=ARATC

BASE COUNT 155 a 118 c 97 g 142 t  
ORIGIN

Query Match 100.0%; Score 263; DB 10; Length 512;  
Best Local Similarity 100.0%; Pred. No. 5e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACATTAACAGATTATTAGCTCATTAACAAAAAAGTCAGAGGTAAGCCATC 60  
DB 18 AAACATTAACAGATTATTAGCTCATTAACAAAAAAGTCAGAGGTAAGCCATC 77  
QY 61 TCAAGCAAGGCTTGATCTGCTACTTAACAAATTCACCAAGACTTCTCTTCCG 120  
DB 78 TCAAGCAAGGCTTGATCTGCTACTTAACAAATTCACCAAGACTTCTCTTCCG 137  
QY 121 TCTCACTCTCCCTTCAGTGTGTCAGCTTACGATTCCTGTCATGCCAAGCC 180  
DB 138 TCTCACTCTCCCTTCAGTGTGTCAGCTTACGATTCCTGTCATGCCAAGCC 197  
QY 181 CAAGTGTGTCATATAAGACCCAGAAATCTACTACTCTTTTCAATTCACAGGGGAA 240  
DB 198 CAAGTGTGTCATATAAGACCCAGAAATCTACTACTCTTTTCAATTCACAGGGGAA 257  
QY 241 TTAACACGCTTCTACCCAGCAT 263  
DB 258 TTAACACGCTTCTACCCAGCAT 280

RESULT 15  
BM676160  
LOCUS BM676160 522 bp mRNA linear EST 27-FEB-2002  
DEFINITION U1-E-EJ0-ahf-b-11-0-U1.s1 U1-E-EJ0 Homo sapiens cDNA clone  
ACCESSION U1-E-EJ0-ahf-b-11-0-U1.3, mRNA sequence.  
VERSION BM676160.1 GI:18986056  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 522)  
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA sequence: 1-28, >AT-richlow-complexity (matched complement)  
Seq primer: M13 Forward  
POLYA=yes

FEATURES  
source  
Location/Qualifiers  
1..522  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="U1-E-EJ0-ahf-b-11-0-U1"  
/clone\_lib="U1-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Paveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="origab: eye; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; U1-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:797-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, ACATTCAGCA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina Paveal and Macular, GTCC; RPE and Choroid, ACCCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG\_LIB=U1-E-EJ0  
TAG\_TISSUE=human retina  
TAG\_SEQ=CCGCG"

BASE COUNT 154 a 124 c 99 g 144 t 1 others  
ORIGIN

Query Match 100.0%; Score 263; DB 14; Length 522;  
Best Local Similarity 100.0%; Pred. No. 5.1e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACATTAACAGATTATTAGCTCATTAACAAAAAAGTCAGAGGTAAGCCATC 60  
DB 19 AAACATTAACAGATTATTAGCTCATTAACAAAAAAGTCAGAGGTAAGCCATC 78  
QY 61 TCAAGCAAGGCTTGATCTGCTACTTAACAAATTCACCAAGACTTCTCTTCCG 120  
DB 79 TCAAGCAAGGCTTGATCTGCTACTTAACAAATTCACCAAGACTTCTCTTCCG 138  
QY 121 TCTCACTCTCCCTTCAGTGTGTCAGCTTACGATTCCTGTCATGCCAAGCC 180

Accession	Sequence	Position
Db	TCTCAGCTCCCTTCAGTGGTGCAGCTTCAGCTATCTCTCGTGTATGATCCCAAGGCC	139
Oy	CAAGTGGTCATCATTAAGACCCAGAGAAATCTACTACCTTTTTCACATTTCAACAGGGGAA	181
Db	CAGGTGGTCATCATTAAGACCCAGAGAAATCTACTACCTTTTTCACATTTCAACAGGGGAA	199
Oy	TTAAACAGCTCTACTCCAGCAT	241
Db	TTAAACAGCTCTACTCCAGCAT	259

Search completed: June 21, 2003, 03:37:16  
Job time : 788.299 secs

Initial incubation: 95 degrees C for 10 minutes

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds

Polymerization: 72 degrees C for 23 seconds

PCR cycles: 30  
Thermal cycler: Perkin Elmer 9700

Protocol:

Primer: each 1  $\mu$ M

AmpliTaq Gold Polymerase: 0.07 units/ul

Total Vol: 5 uL

Buffer:

MgCl<sub>2</sub>: 2.5 mM  
KCl: 50 mM

PH: 8.3

with primer pairs:

Prepared with primer pairs derived from AA004887 -- Unigene.  
Location/Qualifiers

FEATURES	Location/Qualifiers
source	1. .451

Source

1.451

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

```

/clone lib="Human"
/map="11"
/ub="A1C1"

```

```

/clone_id=fullan
49. .302
48. 68
STS

```

```
primer_bind 49. .68
primer_bind complement(283. .302)
```

BASE COUNT	-				
ORIGIN	134 a	77 c	97 g	142 t	1 others

Query Match	Score	DB	Length
99.8%	450	11	451

Query Match	99.86%	Score 450	DB 11	Length 721
Best Local Similarity	100.08%	Pred. No. 1.5e-89		
Mismatches 451	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Matches	451;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

GCAGAGGCGCTTCACCTTTTATTTTCAGTTGTACTCATCTGTGCCACCTGTGCACAAATGGAGTTC 60

Db 1 GCAGAGGCTCCACTTTTATTTCAGTTGTACTCATCTGTGCCACTGTGCCAATGGAGTC 60

61 ACACGCTACCTCAATTCTGAGAGCGCTGCGCAAGNAAAGAGAAAAAGATGCCAGACAGTC 120

Db 61 ACACGCTCACTCAATTCTGAGAGGCTTGGCAAGNAAAGAGAAAAAGATGCCCGACAGTC 120

121 TGTTAGCTGCATTCCTGAGACTAATATCTTTACAGCTGTGAGAAATCACTGCAGGGTT 180

121 TGTAGAGTTGCATCTTCAGACATATATCTTTACAGCTCTTGAGAAATCACTGTCAGGGTT 180

Db 121 TGTAGAGTTCACATCTCAGCAGCAATATCTTTATCAGCTCTTGGAGATCTCTCATGCGGCT 208

181 TATTAAATGCAGATTTTGAGAGATAATTTACGACATAATTTTTTATVAAACTAT 240

Db 181 TATTAAATGCAGATTTTGAAGCATAAATTTTACGACTAATTTTTTTAATAACTAT 240

241 GCAGCATTTGTATTAGAAATTTGCCAAATTTAGAGCTTCAGCATGGAATAATTGC 300

Db 241 GCAGGATTGTTATTAGAGATTGCCAAATTACAGTCTTCAGCATGGAATAATTGG 300

301 CCTTCTGTACAGTCTTCTGTTATAGTGGTAAGAAAGTTTCTTCCAGAAAT 360

Db 301 CCTCTTGTCACAGCTCTTCGTTTATAGTGGGTAAAGAAAGTTCTTCCAGAAAT 360

361 ACCAGGAGAAATCCGATGGTCTTGATAGGAGTTAAATTGTGGAGATGTCACAGACAGCA 420

361 ACACGAGAAATCCGATGCTTCTGATAGGAGTAAATGTGGAGATGTCCACGACACCA 420

Db 361 ACAGCAGAAATCCGATGGTCTGATAGGAGTTAATTGTGGAGATGTGCCAGAGACAGCA 420

QY 421 GCTTCGTGATGGTGACCCACAATGCTGT 451

Db 421 GCCTCGATGGTGACACCAATGCTGT 451

RESULT 2	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT	
642509	SHGC-58593	Human Homo sapiens STS genomic, sequence tagged site.	642509	642509.1	GI:4062074	STS.	Homo sapiens.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 451)	Myers, R.M.	Human STS (1998)	Unpublished (1998)
<p>Contact: Richard M. Myers  Stanford Human Genome Center (SHGC)  Stanford University School of Medicine  Department of Genetics, M-344, Stanford, CA 94305, USA  Tel: 4157259687  Fax: 4157259689  Email: myers@hgc.stanford.edu  Primer A: GCAAAATGAGTCACACGTC  Primer B: GGCCATTATTTCATCGCT  STS size: 254</p>												

## ALIGNMENTS

for SEQ ID NO: 164

RESULT 1  
ABL67130  
ID ABL67130 standard; DNA: 451 BP.  
XX  
AC ABL67130;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5467.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.





for SET ID NO: 164

RESULT 1  
AA004887  
LOCUS  
DEFINITION 451 bp mRNA linear EST 23-JUL-1996  
2h90g01.s1 Soares fetal\_liver\_spleen\_INFLS\_S1 Homo sapiens cDNA

ACCESSION  
AA004887  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

human.

REFERENCE  
AUTHORS  
Mammalia: Eutheria; Primates: Catarrhini; Homiidae: Homo.  
1 (bases 1 to 451)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Tivanskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.

TITLE  
JOURNAL  
COMMENT  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: mob.REGA+ET  
High quality sequence stop: 429.

Mon Jun 23 10:00:35 2003

us-09-964-8:

## FEATURES

Location/Qualifiers

1..451  
/organism="Homo sapiens"  
/db\_xref="GDB:1328361"  
/db\_xref="taxon:9606"  
/clone="IMAGE:428592"  
/clone\_lib="Soares\_fetal\_liver\_spleen\_INFLS\_S1"  
/sex="male"  
/dev\_stage="20 week post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
this is a subtracted version of the original Soares fetal  
liver spleen INFLS library. 1st strand cDNA was primed  
with a Pac I - 0190(6F) primer [5',  
AACTGAGAGATTAAATTAAGATCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Fatima Bonaldo."

## BASE COUNT

134 a 77 c 97 g 142 t 1 others

Query Match 99.8%; Score 450; DB 9; Length 451;  
Best Local Similarity 100.0%; Pred. No. 7,1e-108;  
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	GCAGAGGCGCTCCACTTTTATTTGAGTGTGACATCTGTCCTGCACTGTGCAAAATGAGATC	60
DB	1	GCAGAGGCGCTCCACTTTTATTTGAGTGTGACATCTGTCCTGCACTGTGCAAAATGAGATC	60
OY	61	ACAGCGCTCACTCAATTCGAGAGGCGCTGCAAGNAAAGAAAGATGCCAGAGCAGTC	120
DB	61	ACAGCGCTCACTCAATTCGAGAGGCGCTGCAAGNAAAGAAAGATGCCAGAGCAGTC	120
OY	121	TGTTAGAGTGTGATCTGAGACTAATATCTTACAGCTTGTGAAATCAGTCAAGGTT	180
DB	121	TGTTAGAGTGTGATCTGAGACTAATATCTTACAGCTTGTGAAATCAGTCAAGGTT	180
OY	181	TATTTAAATGAGATTTTGAAGATTAATTTTACGACTAATTTTAAATTAATCTAT	240
DB	181	TATTTAAATGAGATTTTGAAGATTAATTTTACGACTAATTTTAAATTAATCTAT	240
OY	241	GCAGATTTGTTATTTGAGATTTGCCAATTTAGAGCTGTCGAGATGGAATTAATG	300
DB	241	GCAGATTTGTTATTTGAGATTTGCCAATTTAGAGCTGTCGAGATGGAATTAATG	300
OY	301	CCTTCCTGACAGATCTTCTGTTATTAAGTGGTAAAGAAAGTTTCTTCCAGAAAAAT	360
DB	301	CCTTCCTGACAGATCTTCTGTTATTAAGTGGTAAAGAAAGTTTCTTCCAGAAAAAT	360
OY	361	ACAGCAGAAAAATCCGATGTTCTGATAGAGATTATTTGAGAGTGTGCCAGAGACAGCA	420
DB	361	ACAGCAGAAAAATCCGATGTTCTGATAGAGATTATTTGAGAGTGTGCCAGAGACAGCA	420
OY	421	GCTTCGTGATGTGACACCAATGTCGTG	451
DB	421	GCTTCGTGATGTGACACCAATGTCGTG	451

## ALIGNMENTS

for SEQ ID NO: 174

RESULT 1  
ABL67140  
ID ABL67140 standard; DNA; 272 BP.  
XX  
AC ABL67140;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5477.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.



6-552 10 ~0:174

Mon Jun 23 10:00:36 2003

us-09-964

RESULT 1  
LOCUS AA281006 272 bp mRNA linear EST 14-AUG-1997  
DEFINITION 259406.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:705106 3',  
mRNA sequence.  
AA281006  
ACCESSION AA281006.1 GI:1923895  
VERSION AA281006  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 272)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 845 Std Error: 0.00  
Seq primer: -41ml3 fwd. RT from Amerisham  
High quality sequence stop: 252.  
FEATURES  
Source Location/Qualifiers  
1..272  
/organism="Homo sapiens"  
/db\_xref="GDB:5854830"  
/db\_xref="taxon:9606"

/clone="IMAGE:705106"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marcil (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCCTCATTTTCTTTTCTT-3'  
]. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p773 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT	45 a	57 c	47 g	123 t
ORIGIN				
Query Match	100.0%	Score 272;	DB 9;	Length 272;
Best Local Similarity	100.0%	Pred. No. 3e-45;		
Matches 272;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTTTTTTTTAATAGCTGTTCTGTCATATAGTTTATTCCTTATCTTTTGAAC	60	
DB	1	TTTTTTTTTAATAGCTGTTCTGTCATATAGTTTATTCCTTATCTTTTGAAC	60	
QY	61	ATTTATACACCTTATTCATGTTCCCTTTAGATCACTATCTCTTACTCTCTGG	120	
DB	61	ATTTATACACCTTATTCATGTTCCCTTTAGATCACTATCTCTTACTCTCTGG	120	
QY	121	GCTTGAATCCTTGTCTGTTCTGATCTGCTGCTCTTGGGATACCGGGAGTTT	180	
DB	121	GCTTGAATCCTTGTCTGTTCTGATCTGCTGCTCTTGGGATACCGGGAGTTT	180	
QY	181	CCTCTGACCTGCTCTCAGTAGAGAAATGATTTTCATGAGATCTGTTCCCTGATG	240	
DB	181	CCTCTGACCTGCTCTCAGTAGAGAAATGATTTTCATGAGATCTGTTCCCTGATG	240	
QY	241	AGGACGCTGCTCTGCGGAGAAATGCTCTGT	272	
DB	241	AGGACGCTGCTCTGCGGAGAAATGCTCTGT	272	

## ALIGNMENTS

for SEQ ID NO: 180

RESULT 1  
ABL67146  
ID ABL67146 standard; DNA; 319 BP.  
XX  
AC ABL67146;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5483.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

PR	25-SEP-2000	2000US-234923P
PR	25-SEP-2000	2000US-234924P
PR	25-SEP-2000	2000US-235077P
PR	25-SEP-2000	2000US-235082P
PR	25-SEP-2000	2000US-235134P
PR	25-SEP-2000	2000US-235280P
PR	25-SEP-2000	2000US-235637P
PR	26-SEP-2000	2000US-235638P
PR	27-SEP-2000	2000US-235711P
PR	27-SEP-2000	2000US-235720P
PR	27-SEP-2000	2000US-235840P
PR	27-SEP-2000	2000US-235863P
PR	28-SEP-2000	2000US-236028P
PR	28-SEP-2000	2000US-236032P
PR	28-SEP-2000	2000US-236037P
PR	28-SEP-2000	2000US-236038P
PR	28-SEP-2000	2000US-236109P
PR	28-SEP-2000	2000US-236111P
PR	29-SEP-2000	2000US-236842P
PR	29-SEP-2000	2000US-236891P
PR	02-OCT-2000	2000US-237172P
PR	02-OCT-2000	2000US-237173P
PR	02-OCT-2000	2000US-237178P
PR	02-OCT-2000	2000US-237278P
PR	02-OCT-2000	2000US-237294P
PR	02-OCT-2000	2000US-237295P
PR	03-OCT-2000	2000US-237316P
PR	03-OCT-2000	2000US-237425P
PR	03-OCT-2000	2000US-237598P
PR	03-OCT-2000	2000US-237604P
PR	03-OCT-2000	2000US-237606P
PR	03-OCT-2000	2000US-237608P
PR	01-NOV-2000	2000US-244867P
PR	01-NOV-2000	2000US-245084P

PA (AVAL-) AVALON PHARM.  
XX  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,  
PI Soppet DR, Weaver Z;  
XX  
DR WPI: 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set  
XX  
PS Claim 1, SEQ ID 3483; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB616164 to AB670110), or is at least 9% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophagaeal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Malm's tumour.

Query March	100.0%	Score 319;	DB 24;	Length 319;
Best Local Similarity	100.0%;	Pred. No. 1,8e-85;		
Matches 319; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

0Y 1 TTTTTCACCTGCACCATGATTTAAATTTTGTAGTCCCCACCAATGCTAGTCTA 60

Db	1	TTTTTTTCCACGTCACCATGAAATTTAAATTTATGTAGAGGCCCAAAATGGTAGCTA	60
QY	61	TTCGTGTCACATTTGATGACACCACTTCTTATCTCTAAAGATGAGAAATATTTGCT	120
Db	61	TTCCTGTCACATTTGATGAGACCACTTCTTATCTCTAAAGATGAGAAATATTTGCT	120
QY	121	ACTATATATTTTTTTTGGTCATCAACCCAGCCAGATACAAATGCAACTCGATGAAATT	180
Db	121	ACTATATATTTTTTTTGGTCATCAACCCAGCCAGATACAAATGCAACTCGATGAAATT	180
QY	181	CTTAACCATTAATGAGGAGAGGGGCTCCAGGCTAAATGCAAGTATCTTGATTAATGTTTC	240
Db	181	CTTAACCATTAATGAGGAGAGGGGCTCCAGGCTAAATGCAAGTATCTTGATTAATGTTTC	240
QY	241	CCACACACTGGGAGATACACCTCCCCGGCTCCCTCGAAGCTTCCCAAAAGGTGGGGGGG	300
Db	241	CCACACACTGGGAGATACACCTCCCCGGCTCCCTCGAAGCTTCCCAAAAGGTGGGGGGG	300
QY	301	AAGCAGAGAAAAAAAAGG 319	
Db	301	AAGCAGAGAAAAAAAAGG 319	

301 AAGCAGGAGAAAAAAGG 319  
 11111111111111111111  
 301 AAGCAGGAGAAAAAAGG 319

## ALIGNMENTS

for seq ID no 180

RESULT 1  
AA490819  
LOCUS  
DEFINITION aa490819 319 bp mRNA linear EST 15-AUG-1997  
aa49f05.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:824289 3',  
mRNA sequence.  
ACCESSION AA490819  
VERSION AA490819.1 GI:2219992  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 319)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -41m13 fwd. ET from Amersham.



## FEATURES

Source Location/Qualifiers  
 1. 319  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:824289"  
 /clone\_lib="NCI\_CGAP\_GCB1"  
 /tissue\_type="germinal center B cell"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."  
 95 a 73 c 57 g 94 t

## BASE COUNT

Query Match 100.0%; Score 319; DB 9; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-71;  
 Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTTTCCTGTCACCATGAATTTAAATTTATTTAGTGGCCACAAATGCTAGTCTA 60  
 Db 1 TTTTTCCTGTCACCATGAATTTAAATTTATTTAGTGGCCACAAATGCTAGTCTA 60  
 QY 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120  
 Db 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120  
 QY 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGATACAAATGGAACCTCTATGAATATT 180  
 Db 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGATACAAATGGAACCTCTATGAATATT 180  
 QY 181 CTAAGCATATTAATGAGAGGGGCTCCAGGCTAAATGCAAGTATCTTGAATATTTTTC 240  
 Db 181 CTAAGCATATTAATGAGAGGGGCTCCAGGCTAAATGCAAGTATCTTGAATATTTTTC 240  
 QY 241 CCCACCTAGGAATACCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 300  
 Db 241 CCCACCTAGGAATACCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 300  
 QY 301 AAGCAGGAGAAAAAAGG 319  
 Db 301 AAGCAGGAGAAAAAAGG 319

## RESULT 2

AA490870  
 LOCUS  
 DEFINITION  
 aa48b07.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:824149 3', mRNA linear EST 15-AUG-1997  
 mRNA sequence.  
 AA490870  
 VERSION  
 AA490870.1 GI:2220043  
 KEYWORDS  
 EST.  
 SOURCE  
 human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 369)

## REFERENCE

AUTHORS  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL  
 Unpublished (1997)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-ref@mail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/dbp/image/image.html  
 Seq primer: -41m3 fwd. ET from Amersham  
 High quality sequence stop: 327.

## FEATURES

Location/Qualifiers  
 1. 369  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:824149"  
 /clone\_lib="NCI\_CGAP\_GCB1"  
 /tissue\_type="germinal center B cell"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."  
 116 a 83 c 67 g 103 t

## BASE COUNT

Query Match 100.0%; Score 319; DB 9; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-71;  
 Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTTTCCTGTCACCATGAATTTAAATTTATTTAGTGGCCACAAATGCTAGTCTA 60  
 Db 1 TTTTTCCTGTCACCATGAATTTAAATTTATTTAGTGGCCACAAATGCTAGTCTA 60  
 QY 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120  
 Db 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120  
 QY 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGATACAAATGGAACCTCTATGAATATT 180  
 Db 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGATACAAATGGAACCTCTATGAATATT 180  
 QY 181 CTAAGCATATTAATGAGAGGGGCTCCAGGCTAAATGCAAGTATCTTGAATATTTTTC 240  
 Db 181 CTAAGCATATTAATGAGAGGGGCTCCAGGCTAAATGCAAGTATCTTGAATATTTTTC 240  
 QY 241 CCCACCTAGGAATACCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 300  
 Db 241 CCCACCTAGGAATACCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 300  
 QY 301 AAGCAGGAGAAAAAAGG 319  
 Db 301 AAGCAGGAGAAAAAAGG 319

## ALIGNMENTS

for seq ID no: 191

RESULT 1  
ABL67157  
ID ABL67157 standard; DNA; 441 BP.  
XX  
AC ABL67157;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5494.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237279P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrihan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 XX WPI: 2002-188264/24.  
 DR  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set -  
 XX  
 PS Claim 1: SEQ ID 5494; 44pp; English.  
 XX  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytosstatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilm's tumour.  
 CC  
 XX  
 SQ Sequence 441 BP; 158 A; 68 C; 72 G; 143 T; 0 other;

Query Match 100.0%; Score 441; DB 24; Length 441;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-85;  
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATTATTATTAAGCTGAATTTTATTACTAAATTAATCTATGTCACAAAAAATTCGTGCC 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db	1	CATTATTATTAAGCTGAATTTTATTACTAAATTAATCTATGTCACAAAAAATTCGTGCC	60
Qy	61	TGGCGTGAATTTTCACTCATCAAGTGTATACATGATTTTTCATTTCATTACAAAGCAG	120
Db	61	TGGCGTGAATTTTCACTCATCAAGTGTATACATGATTTTTCATTTCATTACAAAGCAG	120
Qy	121	GAGATGAATGTAGACAAAGTGTAGGAACATGCAATTAATTAATTAATTTACAA	180
Db	121	GAGATGAATGTAGACAAAGTGTAGGAACATGCAATTAATTAATTAATTTACAA	180
Qy	181	AAGCAAAAAAATTAACAGTGTACACATATTACTGAGATATATAATTAAGCAACAAC	240
Db	181	AAGCAAAAAAATTAACAGTGTACACATATTACTGAGATATATAATTAAGCAACAAC	240
Qy	241	AATCACAATTAATACAAAGTATTTCTGTGTGTAGTGAAGATACCTATGTGACATC	300
Db	241	AATCACAATTAATACAAAGTATTTCTGTGTGTAGTGAAGATACCTATGTGACATC	300
Qy	301	ATTCAACAAAAAAGTTCCTAATGAATGACATATTGGAAATCATATGATCTCAGCG	360
Db	301	ATTCAACAAAAAAGTTCCTAATGAATGACATATTGGAAATCATATGATCTCAGCG	360
Qy	361	GTTTAATCATTTAGGCTATTACCGTTCCCTTTTGTAGTGAAGTATATCCCAAGTGCA	420
Db	361	GTTTAATCATTTAGGCTATTACCGTTCCCTTTTGTAGTGAAGTATATCCCAAGTGCA	420
Qy	421	GATACGTCTCCAGGTGTAG 441	
Db	421	GATACGTCTCCAGGTGTAG 441	

	RESULT 1			
	LOCUS	N71063		
DEFINITION				
	N71063	441 bp	mRNA	linear
	Z886611.s1 Soares_fetal_lung_NbHl19W	Homo sapiens	cDNA clone	
ACCESSION	IMAGE:2994420.3,			
VERSION	N71063			
KEYWORDS	N71063.1 GI:1227643			
SOURCE	EST.			
ORGANISM	human.			
	Homo sapiens			
REFERENCE	Bukacynska; Metzger: Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 441)			
	Hillier L., Clark N., Dubucque T., Elliston K., Hawkins M., Holman			
	M., Hultman M., Kucaba T., Le M., Lennon G., Marra M., Parsons V.,			
	Rifkin L., Rohlfing T., Soares M., Tan F., Trevaskis E., Waterston			
	R., Williamson A., Wohlmann P. and Wilson R.			
TITLE	The WashU-Merck EST Project			
JOURNAL	Unpublished (1995)			
COMMENT	Contact: Wilson RK			
	Washington University School of Medicine			
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			
	Tel: 314 286 1800			
	Fax: 314 286 1810			
	Email: est@watson.wustl.edu			
	This clone is available royalty-free through LLNL ; contact the			
	IMAGE Consortium (info@image.llnl.gov) for further information.			
	Seq primer: m13 -40 forward			
	High quality sequence stop: 344.			

us-09-96

FEATURES					
	source	Location/Qualifiers			
		1..441			
		/organism="Homo sapiens"			
		/db_xref="GDB:1244344"			
		/db_xref="taxon:9606"			
		/clone="IMAGE:299420"			
		/clone_lib="Soares_fetal_lung_NBHL19w"			
		/dev_stage="19 weeks"			
		/lab_host="DH10B (ampicillin resistant)"			
		/note="Organ: Lung; Vector: pRTV3 (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-GTTCACCAATCTGAAGGGAGCGGCCGCACAATTGTTTGTGTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRTV3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library. Soares fetal heart NBHL19w."			
BASE COUNT	158 a	68 c	72 g	143 t	
ORIGIN					
Query Match	100.0%; Score 441;	DB 14;	Length 441;		
Best Local Similarity	100.0%; Pred. No. 2.2e-75;				
Matches 441;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY	1	CATTCTTTAAGCTGAATTTTATTTACTATAATTAATCATGCAAAAAAAATCTGTGCC	60		
Db	1	CATTATTTAAAGCTGAATTTTATTTTACTATAATTAATCATGCAAAAAAAATCTGTGCC	60		
OY	61	TGGCGTGAAATTCACCTCATTCAAGTGTCACATGATTTTTTCAATTTTCAATTACCAACAG	120		
Db	61	TGGCGTGAAATTCACCTCATTCAAGTGTCACATGATTTTTTCAATTTTCAATTACCAACAG	120		
OY	121	GAGATGTAATGTAGCACAGTGTGTAGAACAATGCAATTAATTAATTAATTTACAA	180		
Db	121	GAGATGTAATGTAGCACAGTGTGTAGAACAATGCAATTAATTAATTAATTTACAA	180		
OY	181	AAGCAAAAATTTAAGAGTAGACACATTTTAACTGATTAATTAATTAATAGACACACT	240		
Db	181	AAGCAAAAATTTAAGAGTAGACACATTTTAACTGATTAATTAATTAATTAATAGACACACT	240		
OY	241	AATCACAATTAATACAAAGTAAATTCCTGTGTTCTAGTACAGATACCCTATGTGACATTC	300		
Db	241	AATCACAATTAATACAAAGTAAATTCCTGTGTTCTAGTACAGATACCCTATGTGACATTC	300		
OY	301	ATTCAACAAAAAAGTCCCTAATGAAATGGACATATTGGGAAATCATATGATATCTCACGG	360		
Db	301	ATTCAACAAAAAAGTCCCTAATGAAATGGACATATTGGGAAATCATATGATATCTCACGG	360		
OY	361	GGTTAAATCATAGGTAACATTTAACCGTTTCCCTTTTAGAGACATTATCCAGTGAGCA	420		
Db	361	GGTTAAATCATAGGTAACATTTAACCGTTTCCCTTTTAGAGACATTATCCAGTGAGCA	420		
OY	421	GATACGTCCCAAGGTGAAG 441			
Db	421	GATACGTCCCAAGGTGAAG 441			

## ALIGNMENTS

RESULT 1  
ABL67165  
ID ABL67165 standard; DNA; 448 BP.  
XX  
AC ABL67165;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5502.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

6- SEQ ID NO:199



US-09-964.

Location/Qualifiers  
1. .448

```

/clone_lib="Soares.NHMPu.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/notes="Organ: mixed (see below); Vector: empty; Res:
"
```

BASE COUNT	127	a	89	c	82	g	150	t
ORIGIN								

Query Match	100.0%;	Score 448;	DB 9;	Length 448;
Best Local Similarity	100.0%;	Pred. No. 5.3e-73;		
Matches 448; Conservative	0;	Mismatches	0;	Gaps 0;

[illegible]

## ALIGNMENTS

RESULT 1  
ABL67310  
ID ABL67310 standard; DNA; 283 BP.  
XX  
AC ABL67310;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5647.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
FN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

For SEQ ID NO:344





96-60-us

```

FEATURES
Source
Location/Qualifiers
1.283
/organism="Homo sapiens"
/db_xref="GDB:5976674"
/db_xref="taxon:9606"
/clone="IMAGE:753807"
/clone_1ib="Soares.NbHMpU.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below): Vector: pT713D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHMp, and fetal heart NbHM19W) were mixed, and as circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of T.M.A.C.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT
63 a 95 c 90 g 35 t
ORIGIN

```

PT involves detecting the level of expression of two or more genes in a liver tissue sample

Claim 1: SEQ ID NO 2679; 298pp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN9203-ABN97455 in a tissue sample. The method of the invention has hepatotrophic, and cytosolic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences.

XX Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;

Query Match 100.0%; Score 165; DB 24; Length 165;  
Best Local Similarity 100.0%; Pred. No. 3,7e-27;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATAGCATTAGACATACATTAATGA 60  
DB 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATAGCATTAGACATACATTAATGA 60  
DB 61 CACACCACTGTTGACAAATGAAAAAAGCAATTCATATTTTCCAGCTTTTAAGTT 120  
QY 61 CACACCACTGTTGACAAATGAAAAAAGCAATTCATATTTTCCAGCTTTTAAGTT 120  
DB 61 CACACCACTGTTGACAAATGAAAAAAGCAATTCATATTTTCCAGCTTTTAAGTT 120  
QY 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTTAA 165  
DB 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTTAA 165

RESULT 2  
ABL62688  
ID ABL62688 standard; DNA; 165 BP.

XX ABL62688;

DT 15-MAY-2002 (first entry)

DE Colon adenocarcinoma related gene sequence SEQ ID NO:1025.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KM cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
KM gene; ds.

OS Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.

PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 25-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 26-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 29-SEP-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.

(AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
Soppet DR, Weaver Z;

WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set

Claim 1: SEQ ID 1025; 44pp; English.

XX The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (1) of a signature gene set, where (1) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (1) has cytosolic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;

Query Match 100.0%; Score 165; DB 24; Length 165;  
Best Local Similarity 100.0%; Pred. No. 3,7e-27;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATAGCATTAGACATACATTAATGA 60  
DB 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATAGCATTAGACATACATTAATGA 60

```

OY      61 CACACCACTTTGACATGAAAAAAGACATTGATTTTCCAGCTTTTAACTT 120
        |||||||
DB      61 CACACCACTTTGACATGAAAAAAGACATTGATTTTCCAGCTTTTAACTT 120
OY      121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTAG 165
        |||||||
DB      121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTTAA 165

RESULT 3
ABL67347
ID      ABL67347 standard: DNA: 165 BP.
XX
XX      ABL67347;
AC
XX      15-MAY-2002 (first entry)
DT
XX      Thyroid cancer related gene sequence SEQ ID NO:5684.
DE
XX      Human; Cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW      stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW      cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX      gene; ds.
XX
OS      Homo sapiens.
XX
XX      WO200194629-A2.
XX
XX      13-DEC-2001.
PD
XX
XX      30-MAY-2001: 2001WO-US10838.
PF
XX
XX      05-JUN-2000; 2000US-209473P.
PR      05-JUN-2000; 2000US-209531P.
PR      18-SEP-2000; 2000US-23133P.
PR      18-SEP-2000; 2000US-233617P.
PR      20-SEP-2000; 2000US-234009P.
PR      20-SEP-2000; 2000US-234034P.
PR      20-SEP-2000; 2000US-234052P.
PR      22-SEP-2000; 2000US-234509P.
PR      22-SEP-2000; 2000US-234567P.
PR      25-SEP-2000; 2000US-234923P.
PR      25-SEP-2000; 2000US-234924P.
PR      25-SEP-2000; 2000US-235072P.
PR      25-SEP-2000; 2000US-235082P.
PR      25-SEP-2000; 2000US-235134P.
PR      25-SEP-2000; 2000US-235280P.
PR      26-SEP-2000; 2000US-235637P.
PR      27-SEP-2000; 2000US-235638P.
PR      27-SEP-2000; 2000US-235711P.
PR      27-SEP-2000; 2000US-235720P.
PR      27-SEP-2000; 2000US-235840P.
PR      27-SEP-2000; 2000US-235863P.
PR      28-SEP-2000; 2000US-236028P.
PR      28-SEP-2000; 2000US-236032P.
PR      28-SEP-2000; 2000US-236033P.
PR      28-SEP-2000; 2000US-236034P.
PR      28-SEP-2000; 2000US-236109P.
PR      28-SEP-2000; 2000US-236111P.
PR      29-SEP-2000; 2000US-236842P.
PR      29-SEP-2000; 2000US-236891P.
PR      02-OCT-2000; 2000US-237172P.
PR      02-OCT-2000; 2000US-237173P.
PR      02-OCT-2000; 2000US-237278P.
PR      02-OCT-2000; 2000US-237294P.
PR      02-OCT-2000; 2000US-237295P.
PR      02-OCT-2000; 2000US-237316P.
PR      03-OCT-2000; 2000US-237425P.
PR      03-OCT-2000; 2000US-237598P.
PR      03-OCT-2000; 2000US-237604P.
PR      03-OCT-2000; 2000US-237606P.
PR      03-OCT-2000; 2000US-237608P.

```

```

PR      01-NOV-2000; 2000US-244867P.
PR      01-NOV-2000; 2000US-245084P.
XX
XX      (AVAL-) AVALON PHARM.
XX
XX      Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI      Soppet DR, Weaver Z;
XX      WPI: 2002-188264/24.
DR
XX
XX      Screening for anti-neoplastic agent involves exposing cells to a
PI      chemical agent to be tested for anti-neoplastic activity, and
PI      determining a change in expression of a gene of a signature gene set
XX
XX      Claim 1; SEQ ID 5684; 44pp; English.
PS
XX
XX      The present invention describes a method (M1) for screening for an
CC      anti-neoplastic agent. The method involves exposing cells to a chemical
CC      agent to be tested for anti-neoplastic activity, determining a change in
CC      expression of at least one gene (I) of a signature gene set, where (I)
CC      comprises a sequence (S) selected from 847 sequences (given in ABL61664
CC      to ABL70110), or is at least 95% identical to (S), where a change in
CC      expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC      activity and can be used in gene therapy. M1 can be used for screening
CC      an anti-neoplastic agent, and can be used for producing a product which
CC      is the data collected with respect to the anti-neoplastic agent as a
CC      result of M1, and the data is sufficient to convey the chemical
CC      structure and/or properties of the agent. M1 can be used in the
CC      treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC      oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC      adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC      infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC      carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX      Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;
SQ

Query Match      100.0%; Score 165; DB 24; Length 165;
Best Local Similarity 100.0%; Pred. No. 3.7e-27;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 CTCTTGAGTAAGTTATTTTGGAGAGCTTCATAGCATTTAGAACATACATAAAATGA 60
        |||||||
DB      1 CTCTTGAGTAAGTTATTTTGGAGAGCTTCATAGCATTTAGAACATACATAAAATGA 60
OY      61 CACACCACTTTGACATGAAAAAAGACATTGATTTTCCAGCTTTTAACTT 120
        |||||||
DB      61 CACACCACTTTGACATGAAAAAAGACATTGATTTTCCAGCTTTTAACTT 120
OY      121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTAG 165
        |||||||
DB      121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTTAA 165

RESULT 4
ABL69408
ID      ABL69408 standard: DNA: 165 BP.
XX
XX      ABL69408;
AC
XX      15-MAY-2002 (first entry)
DT
XX      Prostate cancer related gene sequence SEQ ID NO:7745.
DE
XX      Human; Cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW      stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW      cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX      gene; ds.
XX
OS      Homo sapiens.
XX
XX      WO200194629-A2.
XX
XX      13-DEC-2001.

```

XX 30-MAY-2001; 2001WO-US10838.  
PF 05-JUN-2000; 2000US-209473P.  
XX 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 20-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 22-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 25-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 26-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 29-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
XX (AVAL-) AVALON PHARM.  
XX  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,  
PI Soppet DR, Weaver Z;  
XX  
XX WPI: 2002-188264/24.  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set  
XX  
PS Claim 1: SEQ ID 7745: 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytosstatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oophagel, ovarian, kidney, prostate or pancreatic cancer,

CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms' tumour.  
XX  
SQ Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;  
Query Match 100.0%; Score 165; DB 24; Length 165;  
Best local Similarity 100.0%; Pred. No. 3.7e-27;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCCTTGAGTAACTTATTTTGGAGAGTTCATTAAGCATTTAGACATACATAAATGA 60  
Db 1 CTCCTTGAGTAACTTATTTTGGAGAGTTCATTAAGCATTTAGACATACATAAATGA 60  
QY 61 CACACCACGTGTACATGAAAAAAGCAGATTGATATTTTCAGCTTTTAAAGTT 120  
Db 61 CACACCACGTGTACATGAAAAAAGCAGATTGATATTTTCAGCTTTTAAAGTT 120  
QY 121 AAAAAATGATTCAGTTAAAAACAAAAGTTTACATTTTAAAG 165  
Db 121 AAAAAATGATTCAGTTAAAAACAAAAGTTTACATTTTAAAG 165  
RESULT 5  
ID AAF91861/c  
AAAF91861 standard; cDNA; 5330 BP.  
AC AAF91861;  
XX  
XX 22-MAY-2001 (first entry)  
XX  
DE Human secreted protein-encoding gene 4 cDNA clone HDPFY41, SEQ ID NO:14.  
XX  
KW Human; secreted protein; proliferative disorder; cancer; tumour;  
KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KW cardiovascular disorder; sepsis; diabetes; atherosclerosis;  
KW gastrointestinal disorder; angiodysplasia; kidney disorder;  
KW endocrine disorder; infection; pregnancy-related disorder;  
KW cell culture; chemotaxis; food additive;  
KW binding partner identification; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200118022-A1.  
XX  
XX 15-MAR-2001.  
XX  
XX 31-AUG-2000; 2000WO-US24008.  
XX  
XX 03-SEP-1999; 99US-0152315.  
XX  
XX 03-SEP-1999; 99US-0152317.  
XX  
XX (HUMA-) HUMA GENOME SCI INC.  
XX  
XX Ni J, Baker KP, Birse CE, Piscella M, Komatsoulis GA, Rosen CA;  
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Latleur DW;  
PI Moore PA, Shi Y, Wei Y, Florence KA;  
XX  
XX WPI: 2001-203081/20.  
XX  
XX P-PDB: AAB87345.  
XX  
XX Nucleic acid molecules encoding human secreted proteins, used in  
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
PT Parkinson's diseases and cancers  
XX  
PS Claim 1: Page 490-491; 607pp; English.  
XX  
XX AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted  
CC protein genes, and AAB87345-AAB87413 represent the proteins they encode.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 20:42:18 ; Search time 488.851 Seconds  
(without alignments)  
5466.401 Million cell updates/sec

Title: US-09-964-824c-381

Perfect score: 165  
Sequence: 1 cctcttgagtaacttatctt.....caaaagttagatatcttag 165

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estlin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hnc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hnc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_man: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_rtd: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	165	100.0	165	14	N73808	N73808.yz80g07.s1
2	165	100.0	213	9	A1699181	A1699181.tx6e11.x
3	165	100.0	282	14	R50866	R50866.yz62h08.s1
4	165	100.0	302	9	A1910763	A1910763.wy96h11.x
5	165	100.0	303	9	A1969490	A1969490.wz6f01.x
6	165	100.0	336	9	A1351615	A1351615.gr07f12.x

7	165	100.0	381	9	AA043474	AA043474.zk54e09.s
8	165	100.0	412	14	H08164	H08164.y187e01.s1
9	165	100.0	423	9	A1049699	A1049699.an33d09.x
10	165	100.0	423	10	AM592865	AM592865.hg04b07.x
11	165	100.0	444	9	A1041596	A1041596.ox56h06.x
12	165	100.0	450	10	AM072470	AM072470.xa60d07.x
13	165	100.0	451	9	A1573107	A1573107.tr69f01.x
14	165	100.0	459	10	AM770384	AM770384.h180d09.x
15	165	100.0	462	9	AA479302	AA479302.zv21f08.s
16	165	100.0	484	10	AM512118	AM512118.xx70f03.x
17	165	100.0	540	9	A1768486	A1768486.wh2e02.x
18	165	100.0	545	9	A1857643	A1857643.wk95h01.x
19	165	100.0	549	12	BF061430	BF061430.7j53f07.x
20	163	98.8	502	14	N48185	N48185.yz03d10.s1
21	154	93.3	234	10	AM150540	AM150540.xc52h03.x
22	153	93.2	368	9	AA976261	AA976261.oq67a06.s
23	153	92.7	408	9	A1873695	A1873695.wm2e08.x
24	153	92.7	536	12	BG114119	BG114119.602285531
25	152	92.1	390	9	AA281459	AA281459.zs96d03.s
26	152	92.1	459	10	AM444556	AM444556.ui-h-B13-
27	151.4	91.8	387	9	A1915793	A1915793.wg94a07.x
28	151.4	91.8	568	12	BE879247	BE879247.601491650
29	151	91.5	542	9	A1679751	A1679751.tu76c08.x
30	150.4	91.2	505	9	A1693720	A1693720.wd44e08.x
31	149.2	90.4	511	9	AA527097	AA527097.n126e02.s
32	148.8	90.2	188	9	AA743364	AA743364.ny18d01.s
33	143	86.7	169	10	AM613593	AM613593.hb79f09.x
34	140	84.8	393	10	AM183811	AM183811.xj88b10.x
35	140	84.8	698	14	BM997974	BM997974.ui-h-DT1-
36	136.8	82.9	690	12	BG289398	BG289398.602381427
37	136	82.4	358	9	AA909938	AA909938.om12h12.s
38	134.6	81.6	561	10	AM293978	AM293978.ui-h-BW0-
39	112	67.9	888	9	AL561910	AL561910.AL561910
40	104.2	63.2	256	9	AV152678	AV152678.AV152678
41	104.2	63.2	301	10	BB070199	BB070199.BB070199
42	104.2	63.2	363	10	BE335251	BE335251.ug95h07.x
43	104.2	63.2	392	12	BF450263	BF450263.u267e01.x
44	104.2	63.2	425	14	BM899961	BM899961.ui-M-DJ1-
45	104.2	63.2	427	17	AZ630834	AZ630834.1M0484B12

## ALIGNMENTS

RESULT 1  
N73808  
LOCUS N73808 165 bp mRNA linear EST 19-MAR-1996  
DEFINITION yz80g07.s1 Soares,multiple\_sclerosis\_2NDHMS Homo sapiens cDNA  
ACCESSION N73808  
VERSION N73808.1 GI:1231093  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 165)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Matra,M., Parsons,J.,  
Ritkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston  
,R., Williamson,A., Wohlmann,P. and Wilson,R.  
The Washu-Merck EST Project  
Unpublished (1995)  
TITLE JOURNAL  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (infoimage.llnl.gov) for further information.  
Seq primer: ml3 -40 forward  
High quality sequence stop: 144.



```

/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/Note="Organ: whole brain; Vector: Lactid BA; Site: 1: Not
1: Site 2: Hind III; 1st strand cDNA was primed with a Not
1 - oligo(dT) primer 15';
AAGTGGAGAAATTCGGCCGCGCAGCAATTTTCTTTTCTTTT 31;
Double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lactid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      101 a      43 c      40 g      96 t      2 others
ORIGIN
Query Match      100.0%; Score 165; DB 14; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCCTTGAGTAAGTCTTATTTTGGAGAGTTCATAGCAATGAAATACATATAAATGA 60
    |||||||
Db 10 CTCCTTGAGTAAGTCTTATTTTGGAGAGTTCATAGCAATGAAATACATATAAATGA 69
QY 61 CACACCACTGTTGACAATGAAAAAAGACACATTGATATTTTCCAGCTTTTAACTT 120
    |||||||
Db 70 CACACCACTGTTGACAATGAAAAAAGACACATTGATATTTTCCAGCTTTTAACTT 129
QY 121 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTAGT 165
    |||||||
Db 130 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTAGT 174

RESULT 4
AI910763      302 bp      mRNA      linear      EST 20-DEC-1999
LOCUS      wg96h11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2379141 3',
DEFINITION      mRNA sequence.
ACCESSION      AI910763
VERSION      AI910763.1 GI:5630499
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 302)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 1427 Std Error: 0.00
Seq primer: -400p from Gibco
High quality sequence scop: 297.
Location/Qualifiers
1..302
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2379141"
/clone_lib="NCI-CGAP_Kid11"
/lab_host="DH10B"
/Note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site: 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP

```

```

purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 132376-132911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT      113 a      49 c      44 g      96 t
ORIGIN
Query Match      100.0%; Score 165; DB 9; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCCTTGAGTAAGTCTTATTTTGGAGAGTTCATAGCAATGAAATACATATAAATGA 60
    |||||||
Db 4 CTCCTTGAGTAAGTCTTATTTTGGAGAGTTCATAGCAATGAAATACATATAAATGA 63
QY 61 CACACCACTGTTGACAATGAAAAAAGACACATTGATATTTTCCAGCTTTTAACTT 120
    |||||||
Db 64 CACACCACTGTTGACAATGAAAAAAGACACATTGATATTTTCCAGCTTTTAACTT 123
QY 121 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTAGT 165
    |||||||
Db 124 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTAGT 168

RESULT 5
AI969490      303 bp      mRNA      linear      EST 20-OCT-2000
LOCUS      wz67f01.x1 NCI_CGAP_Mell5 Homo sapiens cDNA clone IMAGE:2563125 3',
DEFINITION      mRNA sequence.
ACCESSION      AI969490
VERSION      AI969490.1 GI:5766308
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 303)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium Dyt Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 449 Std Error: 0.00
Seq primer: -400p from Gibco.
Location/Qualifiers
1..303
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2563129"
/clone_lib="NCI-CGAP_Mell5"
/tissue_type="malignant melanoma, metastatic to lymph
node"
/lab_host="DH10B"
/Note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies."
BASE COUNT      114 a      49 c      44 g      96 t
ORIGIN
Query Match      100.0%; Score 165; DB 9; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



QY 1 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 60  
|||||  
Db 4 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 63  
|||||  
QY 61 CACACCACTGTTGCAATGAAAAAAGAGCATTTGATTTTCCAGCTTTTAAGTT 120  
|||||  
Db 64 CACACCACTGTTGCAATGAAAAAAGAGCATTTGATTTTCCAGCTTTTAAGTT 123  
|||||  
QY 121 AAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAA 165  
|||||  
Db 124 AAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAA 168  
|||||

RESULT 6  
LOCUS AA03474  
DEFINITION 356 bp mRNA linear EST 30-DEC-1998  
IMAGE:1940207.3', mRNA sequence.  
ACCESSION A1351615  
VERSION A1351615  
KEYWORDS GI:4088821  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 356)  
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -40UP from Glibco  
High quality sequence stop: 310.  
FEATURES  
location/Qualifiers  
1..356  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1940207"  
/clone\_lib="Soares\_total\_fetus\_Nb2HF8\_9w"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/note="Vector: p7773D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from pooled 8-9 week  
(total) fetus material with a Not I - oligo(dT) primer [5'  
TGTACCACTGTAAGTGGAGCGCGCCCTTAATTTTCTTTTCTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p7773 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 130 a 53 c 51 g 122 t  
ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 356;  
Best Local Similarity 100.0%; Pred. No. 1.2e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7  
LOCUS AA043474  
DEFINITION 381 bp mRNA linear EST 04-SEP-1996  
IMAGE:486664.3', mRNA sequence.  
ACCESSION AA043474  
VERSION AA043474.1  
KEYWORDS GI:1521330  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 381)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,  
Ritkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisakis,E., Waterston  
R., Williamson,A., Wohlmann,P. and Wilson,R.  
The WashU-Merck EST Project  
Unpublished (1995)  
JOURNAL Contact: Wilson RK  
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 301.  
FEATURES  
location/Qualifiers  
1..381  
/organism="Homo sapiens"  
/db\_xref="GDB:3760378"  
/db\_xref="taxon:9606"  
/clone="IMAGE:486664"  
/clone\_lib="Soares\_pregnant\_uterus\_NBHPU"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Organ: uterus; Vector: p7773-Pac; Site\_1: Not I;  
Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5'  
AAGTGAAGATTCGGCGCCCTTTTCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p7773 vector. Library  
went through one round of normalization. Library  
constructed by M. Fatima Bonaldo."

BASE COUNT 142 a 59 c 59 g 121 t  
ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 381;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 60  
|||||  
Db 1 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 60  
|||||  
QY 61 CACACCACTGTTGCAATGAAAAAAGAGCATTTGATTTTCCAGCTTTTAAAGTT 120  
|||||  
Db 61 CACACCACTGTTGCAATGAAAAAAGAGCATTTGATTTTCCAGCTTTTAAAGTT 120  
|||||  
QY 121 AAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAA 165  
|||||  
Db 121 AAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAA 165  
|||||

RESULT 8  
LOCUS H08164  
H08164 412 bp mRNA linear EST 23-JUN-1995

DEFINITION y187e01.s1 Soares infant brain INIB Homo sapiens cDNA clone  
IMAGE:44945 3', mRNA sequence.  
ACCESSION H08164  
VERSION H08164.1 GI:872986  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 412)  
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Harkins, M., Holman, R., Holman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
CONTACT: Wilson RK  
JOURNAL Washington University School of Medicine  
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 337  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: Promega -21ml3  
High quality sequence stop: 337.  
Location/Qualifiers  
1..412  
/organism="Homo sapiens"  
/db\_xref="Gene:417486"  
/db\_xref="taxon:9606"  
/clone="IMAGE:44945"  
/clone\_lib="Soares infant brain INIB"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: whole brain; Vector: latmid BA; Site: 1: Not I; Site: 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACTGACAAATTCGCCGCGAGCAATTTTATTTTATTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the latmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 143 a 65 c 64 g 139 t 1 others  
ORIGIN

Query Match 100.0%; Score 165; DB 14; Length 412;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCATTTAGCAACATACATAAATGA 60  
|||||  
DB 14 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCATTTAGCAACATACATAAATGA 73  
61 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAAGTT 120  
|||||  
DB 74 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAAGTT 133  
121 AAAAAATGATTCAGTTAAAAAAGTTTAAAGTTTAAAGTT 165  
|||||  
DB 134 AAAAAATGATTCAGTTAAAAAAGTTTAAAGTTTAAAGTTTAAAGTT 178

RESULT 9  
LOCUS A1049699 423 bp mRNA linear EST 09-JUL-1998  
DEFINITION an33d09.x1 Gessler Wlms tumor Homo sapiens cDNA clone  
IMAGE:1700465 3', mRNA sequence.  
ACCESSION A1049699

VERSION A1049699.1 GI:3298816  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 423)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kruman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Sepce, M., Tan, F., Theising, B., White, X., Wylie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
Unpublished (1997)  
CONTACT: Wilson RK  
JOURNAL Washington University School of Medicine  
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40ml3 fwd. ET from AmerSham.  
Location/Qualifiers  
1..423  
/organism="Homo sapiens"  
/db\_xref="Gene:1700465"  
/clone="IMAGE:1700465"  
/clone\_lib="Gessler Wlms tumor"  
/sex="pooled (6)"  
/lab\_host="DH10B"  
/note="Vector: pSPOR1; Site: 1: SalI; Site: 2: NotI; RNA was prepared from a pool of 6 anonymous Wlms' tumor RNAs. RNA was prepared by acid-phenol, followed by one round of oligo dt selection. cDNA library preparation was with the BRL/Life Tech. Superscript Plasmid system. An oligo dT NotI primer for first strand synthesis generated ggcgcgcgcct(n) at the 3' end of the clones. A 5' SalI adaptor was used with sequence 5'-gtcgcgcgcgcgcgcg-3'. Resulting cDNAs were size selected (average size 2 kb), NotI digested, and ligated into NotI/SalI-cut pSPOR1. Library was constructed by Dr. Manfred Gessler."

BASE COUNT 155 a 65 c 60 g 143 t  
ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 423;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCATTTAGCAACATACATAAATGA 60  
|||||  
DB 14 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCATTTAGCAACATACATAAATGA 73  
61 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAAGTT 120  
|||||  
DB 74 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAAGTT 133  
121 AAAAAATGATTCAGTTAAAAAAGTTTAAAGTTTAAAGTTTAAAGTT 165  
|||||  
DB 134 AAAAAATGATTCAGTTAAAAAAGTTTAAAGTTTAAAGTTTAAAGTT 178

RESULT 10  
LOCUS AW592865 423 bp mRNA linear EST 22-MAR-2000  
DEFINITION hg04b07.x1 Soares NFL T.GBC.SI Homo sapiens cDNA clone  
IMAGE:2944597 3', mRNA sequence.  
ACCESSION AW592865  
VERSION AW592865.1 GI:7280057  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 423)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
JOURNAL Tumor Gene Index  
COMMENT Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-rt@mail.nih.gov](mailto:cgapbs-rt@mail.nih.gov)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -400P from Glbco  
High quality sequence stop: 209.  
Location/Qualifiers  
1..423  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="2944597"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not 1; Site\_2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBHL19W, testis NHT, and B-cell  
NCI-CGAP-GCB1) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo."  
BASE COUNT 157 a 66 c 63 g 137 t  
ORIGIN  
Query Match 100.0%; Score 165; DB 10; Length 423;  
Best Local Similarity 100.0%; Pred. No. 1,1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 CTCCTTGAGTACTTTATTTGGAGAGGTTCCATAGCATAGAACATACATAAATGA 60  
Db 6 CTCCTTGAGTACTTTATTTGGAGAGGTTCCATAGCATAGAACATACATAAATGA 65  
Oy 61 CACACCACTGTTGCAATGAAAAAACAACGATTTGATTTTCCAGCTTTTAAGTT 120  
Db 66 CACACCACTGTTGCAATGAAAAAACAACGATTTGATTTTCCAGCTTTTAAGTT 125  
Oy 121 AAAAAATGATTCAGTTAAACAAACAAACAGTTAGATATTATTAG 165  
Db 126 AAAAAATGATTCAGTTAAACAAACAAACAGTTAGATATTATTAG 170  
RESULT 11  
LOCUS A1041596 444 bp mRNA linear EST 24-SEP-1998  
DEFINITION 0x56h06.x1 Soares\_NHMPu\_S1 Homo sapiens cDNA clone IMAGE:1660379  
ACCESSION A1041596  
VERSION A1041596.1 GI:3280790  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 444)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
JOURNAL Tumor Gene Index  
COMMENT Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-rt@mail.nih.gov](mailto:cgapbs-rt@mail.nih.gov)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert length: 827 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham

FEATURES  
source  
High quality sequence stop: 414.  
Location/Qualifiers  
1..444  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="1660379"  
/clone\_lib="Soares\_NHMPu\_S1"  
/tissue\_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/note="Organ: mixed (see below); Vector: pT7T3D-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Not 1;  
Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NBH, pregnant uterus  
NBHPV, and fetal heart NBHL19W) were mixed, and ss circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries.  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."  
BASE COUNT 166 a 69 c 71 g 138 t  
ORIGIN  
Query Match 100.0%; Score 165; DB 9; Length 444;  
Best Local Similarity 100.0%; Pred. No. 1,1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 CTCCTTGAGTACTTTATTTGGAGAGGTTCCATAGCATAGAACATACATAAATGA 60  
Db 4 CTCCTTGAGTACTTTATTTGGAGAGGTTCCATAGCATAGAACATACATAAATGA 63  
Oy 61 CACACCACTGTTGCAATGAAAAAACAACGATTTGATTTTCCAGCTTTTAAGTT 120  
Db 64 CACACCACTGTTGCAATGAAAAAACAACGATTTGATTTTCCAGCTTTTAAGTT 123  
Oy 121 AAAAAATGATTCAGTTAAACAAACAAACAGTTAGATATTATTAG 165  
Db 124 AAAAAATGATTCAGTTAAACAAACAAACAGTTAGATATTATTAG 168  
RESULT 12  
LOCUS AM072470 450 bp mRNA linear EST 20-OCT-2000  
DEFINITION xa06d07.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
IMAGE:2567533 3', mRNA sequence.  
ACCESSION AM072470  
VERSION AM072470.1 GI:6027468  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 450)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
JOURNAL Tumor Gene Index  
COMMENT Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-rt@mail.nih.gov](mailto:cgapbs-rt@mail.nih.gov)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert length: 1130 Std Error: 0.00  
Seq primer: -40UP from Glbco  
High quality sequence stop: 447.  
Location/Qualifiers  
1..450  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="2567533"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI-CGAP\_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 170 a 70 c 140 t

ORIGIN

Query Match 100.0%; Score 165; DB 10; Length 450;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTCCTTGAGTAACCTTATTTGGAGAGCTTCATAGCATTAAGCACTACATAAATGA 60  
|||||  
DB 4 CTCCTTGAGTAACCTTATTTGGAGAGCTTCATAGCATTAAGCACTACATAAATGA 63  
61 CACACCACTGTTGACAATGAAAAAAGACATTGATTTCCAGCTTTTAAGTT 120  
|||||  
DB 64 CACACCACTGTTGACAATGAAAAAAGACATTGATTTCCAGCTTTTAAGTT 123  
121 AAAAAATGATTCAGTTAAACAAACAAAGTTTGAATATTTTAC 165  
|||||  
DB 124 AAAAAATGATTCAGTTAAACAAACAAAGTTTGAATATTTTAC 168

RESULT 13  
LOCUS A1573107 451 bp mRNA linear EST 14-MAY-1999  
DEFINITION uc69f01.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2223577 3',  
mRNA sequence.  
ACCESSION A1573107  
VERSION A1573107.1 GI:4536481  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 451)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image.html](http://www.bio.llnl.gov/bbrp/image.html)  
Insert Length: 3468 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 405  
POLYA=NO.

FEATURES  
SOURCE Location/Qualifiers  
1..451  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2223577"  
/clone\_lib="NCI\_CGAP\_Pan1"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.72 kb. Life Technologies catalog #:  
11548-013"

BASE COUNT 169 a 71 c 70 g 141 t

ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 451;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTCCTTGAGTAACCTTATTTGGAGAGCTTCATAGCATTAAGCACTACATAAATGA 60  
|||||  
DB 9 CTCCTTGAGTAACCTTATTTGGAGAGCTTCATAGCATTAAGCACTACATAAATGA 68  
61 CACACCACTGTTGACAATGAAAAAAGACATTGATTTCCAGCTTTTAAGTT 120  
|||||  
DB 69 CACACCACTGTTGACAATGAAAAAAGACATTGATTTCCAGCTTTTAAGTT 128  
121 AAAAAATGATTCAGTTAAACAAACAAAGTTTGAATATTTTAC 165  
|||||  
DB 129 AAAAAATGATTCAGTTAAACAAACAAAGTTTGAATATTTTAC 173

RESULT 14  
LOCUS AW770384 459 bp mRNA linear EST 04-MAY-2000  
DEFINITION h180d09.x1 NCI\_CGAP\_Kid13 Homo sapiens cDNA clone IMAGE:3007505 3',  
mRNA sequence.  
ACCESSION AW770384  
VERSION AW770384.1 GI:7702426  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 459)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life  
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The  
I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
Seq primer: -400P from Gibco  
High quality sequence stop: 415.

FEATURES  
SOURCE Location/Qualifiers  
1..459  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3007505"  
/clone\_lib="NCI\_CGAP\_Kid13"  
/tissue\_type="2 pooled Wilms' tumors, one primary and one  
metastatic to brain"  
/lab\_host="DH10B"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies."

BASE COUNT 173 a 71 c 73 g 142 t

ORIGIN

Query Match 100.0%; Score 165; DB 10; Length 459;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTCCTTGAGTAACCTTATTTGGAGAGCTTCATAGCATTAAGCACTACATAAATGA 60  
|||||  
DB 6 CTCCTTGAGTAACCTTATTTGGAGAGCTTCATAGCATTAAGCACTACATAAATGA 65  
61 CACACCACTGTTGACAATGAAAAAAGACATTGATTTCCAGCTTTTAAGTT 120  
|||||  
DB 66 CACACCACTGTTGACAATGAAAAAAGACATTGATTTCCAGCTTTTAAGTT 125

OY 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTAG 165  
DB 126 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTAG 170

RESULT 15  
AA479302

LOCUS AA479302 462 bp mRNA linear EST 08-AUG-1997  
DEFINITION zv21f08.s1 Soares\_NhHMPu\_s1 Homo sapiens cDNA clone IMAGE:754311  
3', mRNA sequence.

ACCESSION AA479302

VERSION AA479302.1 GI:2207858

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

AUTHORS Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo:

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,

Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie

, T., Waterston, R. and Wilson, R.

Washu-Merck EST Project 1997

Unpublished (1997)

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -41m13 fwd. ET from Amersham

High quality sequence stop: 456.

Location/Qualifiers

1..462

/organism="Homo sapiens"

/db\_xref="GDB:5977197"

/db\_xref="taxon:9606"

/clone="IMAGE:754311"

/tissue="Soares\_NhHMPu\_s1"

/clone\_id="Soares\_NhHMPu\_s1"

/tissue\_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab\_host="DH10B"

/note="Organ: mixed (see below): Vector: pT7T3D-Pac

(Pharmacia) with a modified polylinker; Site\_1: Not I;

Site\_2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NbH, pregnant uterus

NbHpu, and fetal heart NbH19W) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA

was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from pools of

5,000 clones made from the same 3 libraries. The pools

consisted of I.M.A.G.E. clones 260232-265223,

340488-345479, and 484488-489479."

BASE COUNT 176 a 70 c 75 g 141 t

ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 462;

Best Local Similarity 100.0%; Pred. No. 1..1e-21;

Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTCTTTGAGTAACTTTATTTTGGAGAGGTTCCATTAAGCATTTAGAACATATAAATGA 60

DB 1 CTCTTTGAGTAACTTTATTTTGGAGAGGTTCCATTAAGCATTTAGAACATATAAATGA 60

OY 61 CACACCACTGTGACATGAAAAAACAAGCATTTGATATTTTCCAGCTTTTAAAGT 120

DB 61 CACACCACTGTGACATGAAAAAACAAGCATTTGATATTTTCCAGCTTTTAAAGT 120

OY 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTAG 165

DB 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTAG 165

Search completed: June 21, 2003, 03:38:08  
Job time : 496.851 secs

## ALIGNMENTS

A- SEQ ID NO: 390

RESULT 1  
ABL67356  
ID ABL67356 standard; DNA; 386 BP.  
XX  
AC ABL67356;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5693.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

PR	25-SEP-2000	2000US-234923P
PR	25-SEP-2000	2000US-234924P
PR	25-SEP-2000	2000US-235077P
PR	25-SEP-2000	2000US-235082P
PR	25-SEP-2000	2000US-235134P
PR	25-SEP-2000	2000US-235280P
PR	26-SEP-2000	2000US-235637P
PR	26-SEP-2000	2000US-235638P
PR	27-SEP-2000	2000US-235711P
PR	27-SEP-2000	2000US-235720P
PR	27-SEP-2000	2000US-235840P
PR	27-SEP-2000	2000US-235863P
PR	28-SEP-2000	2000US-236028P
PR	28-SEP-2000	2000US-236032P
PR	28-SEP-2000	2000US-236033P
PR	28-SEP-2000	2000US-236034P
PR	28-SEP-2000	2000US-236109P
PR	28-SEP-2000	2000US-236111P
PR	29-SEP-2000	2000US-236842P
PR	29-SEP-2000	2000US-236891P
PR	02-OCT-2000	2000US-237172P
PR	02-OCT-2000	2000US-237173P
PR	02-OCT-2000	2000US-237278P
PR	02-OCT-2000	2000US-237294P
PR	02-OCT-2000	2000US-237295P
PR	02-OCT-2000	2000US-237316P
PR	03-OCT-2000	2000US-237425P
PR	03-OCT-2000	2000US-237598P
PR	03-OCT-2000	2000US-237604P
PR	03-OCT-2000	2000US-237606P
PR	03-OCT-2000	2000US-237608P
PR	01-NOV-2000	2000US-244867P
PR	01-NOV-2000	2000US-245084P

Claim 1; SEQ ID 5693; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

50 Sequence 386 BP; 142 A; 56 C; 63 G; 122 T; 3 other;

Query Match	99.2%	Score 383;	DB 24;	Length 386;
Best Local Similarity	100.0%	Pred. No. 5.8e-68;		
Matches 386;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

1 AAATTATATATACATGTTTATTAGAGCACAACCTTTATGTAAATTTACATTTAATGA 60

Db	AAATPATAATATTCAGCTTTATTAGACACACACTTTTATGTAAAATTACATTTAATGA	60
QY	AAAAAATCAAAAATATTATTACAAAATCTTGGAAAGACAGATGTGCATTTGTTCAATTAACAAAT	120
Db	AAAAAATCAAAAATATTATTACAAAATCTTGGAAAGACAGATGTGCATTTGTTCAATTAACAAAT	120
QY	CCAAAGTAGAAAATTAACAAATCCTTTAAAACTCACATTTATTAGAGTGTGTTACAAATT	180
Db	CCAAAGTAGAAAATTAACAAATCCTTTAAAACTCACATTTATTAGAGTGTGTTACAAATT	180
QY	CTTGTGTTAAAGAGCGAGCTACCAAGTTATACATAATATAAGAAAGACAGCTTGCTA	240
Db	CTTGTGTTAAAGAGCGAGCTACCAAGTTATACATAATATAAGAAAGACAGCTTGCTA	240
QY	GGGTACATTTCCCATTTGAAAATCATCTAGTGGTCTCTTTTACACCATTTAGGGGGATTGTTAA	300
Db	GGGTACATTTCCCATTTGAAAATCATCTAGTGGTCTCTTTTACACCATTTAGGGGGATTGTTAA	300
QY	ATGGGNNAAAAAAATCAATATATACTCATATAGGCGTTCAAAAATTTGGTAACTGTACCCC	360
Db	ATGGGNNAAAAAAATCAATATATACTCATATAGGCGTTCAAAAATTTGGTAACTGTACCCC	360
QY	NATACCTGGGNNATGGAGGCGTGTGC	386
Db	NATACCTGGGNNATGGAGGCGTGTGC	386





```

Query Match          98.6%; Score 348; DB 14; Length 353;

BASE COUNT          107 a          53 c          72 g          116 t          5 others
ORIGIN
Seq primer: Promega -21m13
High quality sequence stop: 310.
Location/Qualifiers
1. .353
/organism="Homo sapiens"
/db_xref="GDB:415859"
/db_xref="taxon:9606"
/clone="IMAGE:43318"
/clone_id="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5';
AAGTGGAGAAATTCGCGCCGACAGAAATTTTTTT 3']";
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lambda BA vector
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaudo."

```

XX	Claim 1; SEQ ID 1584; 44pp; English.
PS	
XX	The present invention describes a method (M1) for screening for an
CC	anti-neoplastic agent. The method involves exposing cells to a chemical
CC	agent to be tested for anti-neoplastic activity, determining a change in
CC	expression of at least one gene (I) of a signature gene set, where (I)
CC	comprises a sequence (S) selected from 847 sequences (given in ABU6164
CC	to AB10110), or is at least 95% identical to (S), where a change in
CC	expression is indicative of anti-neoplastic activity. (I) has cytosolic
CC	activity and can be used in gene therapy. M1 can be used for screening
CC	an anti-neoplastic agent, and can be used for producing a product which
CC	is the data collected with respect to the anti-neoplastic agent as a
CC	result of M1, and the data is sufficient to convey the chemical
CC	structure and/or properties of the agent. M1 can be used in the
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC	oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC	carcinoma, papillary carcinoma and Wilms' tumour.
CC	
SQ	Sequence 353 BP; 107 A; 53 C; 72 G; 116 T; 5 other:
Query Match	98.6%; Score 348; DB 24; Length 353;
Best Local Similarity	100.0%; Pred. No. 1.6e-68;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 TTTTNTTTTTTTTTTTTTTCTGCTCACAATGTCAATTTTATTCACACTAGTGCAAACT 60
DB	1 TTTTNTTTTTTTTTTTTTTCTGCTCACAATAATGTCATTTTATTCACACTAGTGCAAACT 60
OY	61 AAATACATAAATTCGAAGAAGTGAAGTGCATGTTCAAACACAAAATGGGACAAATGATTA 120
DB	61 AAATACATAAATTCGAAGAAGTGAAGTGCATGTTCAAACACAAAATGGGACAAATGATTA 120
OY	121 GAATCGCATGTTTTCACAGGTACTACTATTATTTAAAAAAAATCACAAANNAAGAAA 180
DB	121 GAATCGCATGTTTTCACAGGTACTACTATTATTTAAAAAAAATCACAAANNAAGAAA 180
OY	181 ATGTTATCACACACAAAGTAGGATTTAGCAAGNAGNAAATCTGGGACGTGCTCTAGNA 240
DB	181 ATGTTATCACACACAAAGTAGGATTTAGCAAGNAGNAAATCTGGGACGTGCTCTAGNA 240
OY	241 GGTTTAAAAACATTTTCATGGCATTTTGTGAGTTGCTGTGGAGAAGTTGTTTTTATTTGTCC 300
DB	241 GGTTTAAAAACATTTTCATGGCATTTTGTGAGTTGCTGTGGAGAAGTTGTTTTTATTTGTCC 300
OY	301 ACCGTAATCTGGGCAACATCGGGGGCTTACTCTACACTCTCGGACACTGTGCG 353
DB	301 ACCGTAATCTGGGCAACATCGGGGGCTTACTCTACACTCTCGGACACTGTGCG 353
RESULT 4	
ABL67431	ABL67431 standard; DNA; 353 BP.
XX	
AC	ABL67431;
XX	
DT	15-MAY-2002 (first entry)
DE	Thyroid cancer related gene sequence SEQ ID NO:5768.
XX	
Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;	
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancers;	
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX	gene; ds.
OS	Homo sapiens.
PX	
XN	WO200194629-A2.
XX	
13-DEC-2001.	
XX	

PF	30-MAY-2001;	2001WO-US10838.
XX		
PR	05-JUN-2000;	2000US-209473P.
PR	05-JUN-2000;	2000US-209531P.
PR	18-SEP-2000;	2000US-235133P.
PR	18-SEP-2000;	2000US-235617P.
PR	20-SEP-2000;	2000US-234009P.
PR	20-SEP-2000;	2000US-234034P.
PR	20-SEP-2000;	2000US-234052P.
PR	22-SEP-2000;	2000US-234509P.
PR	22-SEP-2000;	2000US-234567P.
PR	25-SEP-2000;	2000US-234923P.
PR	25-SEP-2000;	2000US-234924P.
PR	25-SEP-2000;	2000US-235077P.
PR	25-SEP-2000;	2000US-235082P.
PR	25-SEP-2000;	2000US-235134P.
PR	25-SEP-2000;	2000US-235280P.
PR	26-SEP-2000;	2000US-235637P.
PR	26-SEP-2000;	2000US-235638P.
PR	27-SEP-2000;	2000US-235711P.
PR	27-SEP-2000;	2000US-235720P.
PR	27-SEP-2000;	2000US-235840P.
PR	27-SEP-2000;	2000US-235863P.
PR	28-SEP-2000;	2000US-236028P.
PR	28-SEP-2000;	2000US-236032P.
PR	28-SEP-2000;	2000US-236033P.
PR	28-SEP-2000;	2000US-236034P.
PR	28-SEP-2000;	2000US-236109P.
PR	28-SEP-2000;	2000US-236111P.
PR	29-SEP-2000;	2000US-236842P.
PR	29-SEP-2000;	2000US-236891P.
PR	02-OCT-2000;	2000US-237172P.
PR	02-OCT-2000;	2000US-237173P.
PR	02-OCT-2000;	2000US-237276P.
PR	02-OCT-2000;	2000US-237294P.
PR	02-OCT-2000;	2000US-237295P.
PR	03-OCT-2000;	2000US-237316P.
PR	03-OCT-2000;	2000US-237425P.
PR	03-OCT-2000;	2000US-237598P.
PR	03-OCT-2000;	2000US-237604P.
PR	03-OCT-2000;	2000US-237606P.
PR	01-NOV-2000;	2000US-237608P.
PR	01-NOV-2000;	2000US-244867P.
PR	01-NOV-2000;	2000US-245084P.

PT Screening for anti-neoplastic agent involves exposing cells to a  
PR chemical agent to be tested for anti-neoplastic activity, and  
XX determining a change in expression of a gene of a signature gene set -  
PV  
PS Claim 1; SEQ ID 5768; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB161664 to AB170110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, lymphocytic leukaemia, carcinoma, clear cell cancer, infiltrating ductal cancer

CC Infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour.  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
Sequence 353 BP; 107 A; 53 C; 72 G; 116 T; 5 other;

Query Match	98.6%;	Score 348;	DB 24;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 1.6e-68;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Oy	1	T T T T T T T T T T T T T T T T T C T C A C A A G T C A A T T T A T G C A C A G T G C C A C A A C T	60
Dd	1	T T T T T T T T T T T T T T T T T G C T C C A C A A G T C A A T T T A T G A C A C T A G T G C C A C A C T	60
Oy	61	A A A T R C A T A A T T G C A A A G A A G T G A A C G T T T A A A C A G A A T G T G A C A T G A G T T A	120
Dd	61	A A A T R C A T A A T T G C A A A G A A G T G A A C G T T T C A A A C A G A A T G T G A C A T G A G T T A	120
Oy	121	G A A C T G C A G T N T T C A A G T A C T A C A C T A T A T T T A	180
Dd	121	G A A C T G C A G T N T T C A A G A G T A C T A C A C T A T A T T T A	180
Oy	181	A T G T T A T A C T R C A A G T A G G G A T T T A G A A A G A A A A A T T C T G G G C A C T C G T C I A S N A	240
Dd	181	A T G T T A T A C P A C A A G T A G G G A T T T A G A A A G A A N A A A T T T G G G C A C T C G T C P A G N A	240
Oy	241	G G G T T A A A A C A T T T C A T G S C A T T T T G A G T G C T G T G A G A G A G T C T T T T T A T T T G T C C	300
Dd	241	G G G T T A A A A C A T T T C A T G S C A T T T T G A G T T C G T G T G A G A G T T G T T T T T A T T T G T C C	300
Oy	301	A C C G T A A C T G G G C A A C A A T C C G G G G G C T T A C C T T C A G C T C C G G C A C T C T G G C	353
Dd	301	A C C G T A A C T G G G C A A C A C T C C G G G G G C T T A C C T T C A G C T C T G G C A C T G T G C G	353